DEFENCE R&D DÉFENSE

The relative impact of generic head-related transfer functions and signal bandwidth on auditory localization:

Implications for the design of three-dimensional audio displays

G.R. Arrabito
Defence and Civil Institute of Environmental Medicine

J.R. Mendelson
Department of Speech-Language Pathology, University of Toronto

DISTRIBUTION STATEMENT A
Approved for Public Release
Distribution Unlimited

Defence R&D Canada

Technical Report

DCIEM TR 2000-067

December 2000

National Defence Défense nationale

Canada

20010116 050

The relative impact of generic head-related transfer functions and signal bandwidth on auditory localization:

Implications for the design of three-dimensional audio displays

G.R. Arrabito
Defence and Civil Institute of Environmental Medicine

J.R. Mendelson
Department of Speech-Language Pathology, University of Toronto

Defence and Civil Institute of Environmental Medicine

Technical Report

DCIEM TR 2000-067

December 2000

Author

G. Robert Arrabito

Approved by

David Beevis

Head, Human Factors of Command Systems Section

Approved for release by

Kathy Sutton

Chair, DCIEM Document Review and Library Committee

Abstract

Virtual auditory technology is being considered to cue armoured vehicle or air crew, via headphones of the communication system, to the spatial locations of potential lethal threats. Auditory localization in virtual auditory space (VAS) on the horizontal plane was investigated in this paper as a function of seven generic head-related transfer functions (i.e., digital filters for synthesizing the location of a sound in VAS), signal bandwidth (low-pass 3 kHz, high-pass 3 kHz and low-pass 14 kHz), and listening environment (quiet and in the presence of diffuse ambient Leopard tank noise). Testing was also conducted in the free-field which partially served to psychoacoustically validate the VAS conditions. The outcome of this preliminary study revealed that subject performance was better in free-field than in VAS. In the latter condition, subject performance was not significantly affected by type of generic head-related transfer function. Localization accuracy using the broadband stimulus was not significantly better than with the low-pass 3 kHz stimulus. Performance in the quiet condition was relatively better than in the noise condition. The implications of these results for implementation of a 3-D audio display into military environments and recommendations for future research are discussed.

Résumé

L'écoute virtuelle est une technologie envisagée pour aider les équipages de blindé et d'aéronef à localiser dans l'espace, à l'aide d'écouteurs de télécommunication, les menaces meurtrières potentielles. Le présent document traite de la localisation auditive dans un espace auditif virtuel (EAV) plan horizontal comme fonction de sept fonctions de transfert génériques, asservies aux mouvements de la tête (représentées par des filtres numériques pour synthétiser la position d'un son dans un EAV), de la bande passante des signaux (passe-bas 3 kHz, passe-haut 3 kHz et passe-bas 14 kHz) et des conditions d'écoute (calme et présence d'un bruit diffus ambiant de char Leopard). Des essais ont aussi été menés en champ libre, en partie pour valider l'EAV sur le plan psychoacoustique. Selon cette étude préliminaire, les sujets ont un meilleur rendement en champ libre qu'en EAV. Dans ce dernier cas, le rendement dépend peu du type de fonction de transfert générique, asservie aux mouvements de la tête. La précision de la localisation en présence du stimulus à large bande n'est pas sensiblement meilleure qu'en présence du stimulus passe-bas de 3 kHz. Le rendement en situation de calme est relativement meilleur qu'en présence de bruit. Nous traitons de l'incidence des résultats sur une présentation audio 3-D dans un contexte militaire et faisons des propositions de recherches futures.

This page intentionally left blank.

Executive summary

A sound source that is presented over headphones can be made to appear as though it originated in the listener's natural free-field environment. The technology used to create this perception is a three-dimensional (3-D) audio display. Digital filters, termed head-related transfer functions (HRTFs), are used to synthesize the location of a sound in virtual auditory space (VAS). In a general-purpose 3-D audio display, localization accuracy may depend on the source positions used, the type of stimuli and HRTFs (personal versus generic), and the localization proficiency and experience of the listeners. Virtual auditory technology is being considered to cue armoured vehicle or air crew to the spatial locations of potential lethal threats. However, there are a number of concerns about how well 3-D audio displays will function in a military environment. These concerns include the suitability of generic HRTFs versus individually tailored ones for localizing sound sources in VAS, the effects of the bandwidth limitations imposed by typical communication systems, and the effects of diffuse ambient noise. This study was a preliminary investigation of these concerns.

In the present study, testing of localization performance in VAS and free-field was assessed in quiet (71 dB(A)), and in the presence of diffuse ambient Leopard tank noise (approximately 110 dB(A)) in VAS. The free-field testing partially served to psychoacoustically validate the VAS conditions. The VAS and free-field speaker array configuration consisted of eight speaker positions, spanning 360° on the horizontal plane around the listener. The acoustic stimulus was white noise band-limited in one of three ways: low-pass 3 kHz, high-pass 3 kHz, and low-pass 14 kHz. The low-pass 3 kHz was chosen to reflect the conservative upper cutoff frequency of the communication system. The contribution of monaural and spectral cues could be observed in the high-pass 3 kHz. The low-pass 14 kHz stimulus allowed an assessment of the interaural differences, monaural, and spectral cues in a combination that would be available in a broader bandwidth. In the free-field condition, additional testing consisted of changing the low- and high-pass 3 kHz cutoff frequencies to low- and high-pass 4 kHz, respectively, in order to determine if the frequency of reversals in free-field could be decreased. Seven generic HRTFs were used.

This preliminary investigation revealed that localization accuracy, as measured by average percent correct and front/back reversals, was higher in free-field compared to the two VAS conditions. Average localization performance in the free-field low- and high-pass 4 kHz cutoff frequency conditions was slightly better than the free-field low- and high-pass 3 kHz cutoff frequency conditions. Given this latter result, it is assumed that this improved level in performance would also be observed in VAS. Subject performance in VAS was not significantly affected by type of generic HRTF; localization accuracy using the broadband stimulus was not significantly better than with the low-pass 3 kHz stimulus. This finding suggests that the role of spectral cues is minimal for sound sources located on the horizontal plane and implies that the restriction of the bandwidth of the communication system to 3.5 kHz might not significantly impede user localization accuracy in VAS. Localization performance was degraded in the presence of diffuse ambient Leopard tank noise relative to the quiet condition suggesting that 3-D audio technology may not yet be very useful in present-day noisy military environments.

The present data are limited to the choice of spatial positions and stimuli. It has been shown that performance in VAS is more accurate and results in fewer localization reversals with personal HRTFs compared to generic ones. However, personal HRTFs are traditionally derived from binaural measurements in the ears of the end-listener seated in an anechoic chamber. This requires a substantial investment in infrastructure and equipment, and is presently impractical in most applications. Further research is required to quickly and accurately select and/or modify a generic HRTF for the targeted application. The effect of diffuse ambient noise on user performance with either personal or generic HRTFs also requires further investigation. The hardware limitation imposed on the communication bandwidth needs to be addressed

particularly when virtual sound sources are presented off the horizontal plane. Until the above issues are more fully understood and resolved it may be prudent to proceed cautiously before the adoption of a 3-D audio system into critical mission applications.

Sommaire

Un son peut être transmis sur casque d'écoute en donnant l'impression qu'il provient du champ libre naturel dans lequel se trouve l'auditeur. La technologie utilisée pour produire cette impression consiste en une présentation audio en trois dimensions (3-D). Des filtres numériques, faisant fonction de transfert asservie aux mouvements de la tête (FTAMT), synthétisent la position d'un son dans un espace auditif virtuel (EAV). Dans une présentation audio 3-D polyvalente, la précision de la localisation peut dépendre de la position de la source utilisée, du type de stimulus et des FTAMT (personnalisées ou génériques), ainsi que de l'adresse et de l'expérience de l'auditeur en matière de localisation. L'écoute virtuelle est une technologie envisagée pour aider les équipages de blindé et d'aéronef à localiser dans l'espace les menaces meurtrières potentielles. L'efficacité d'une présentation audio 3-D pose toutefois plusieurs problèmes dans un contexte militaire. Ces problèmes tiennent à la pertinence des FTAMT, personnalisées ou génériques, pour localiser des sources sonores dans un EAV, aux effets des limites de largeur de bande imposées par les systèmes de télécommunication courants et aux effets du bruit diffus ambiant. La présente étude est une première recherche sur ces problèmes.

Dans la présente étude, nous avons évalué des essais de rendement de localisation dans un EAV et en champ libre, dans des conditions de calme (71 dB(A)), et en présence du bruit diffus ambiant d'un char Leopard (environ 110 dB(A)) dans un EAV. Les essais en champ libre ont servi en partie à valider l'EAV sur le plan psychoacoustique. La disposition des haut-parleurs dans l'EAV et en champ libre était la suivante : huit haut-parleurs répartis sur 360° dans le plan horizontal autour de l'auditeur. Le stimulus acoustique était un bruit blanc sur trois bandes limitées : passe-bas 3 kHz, passe-haut 3 kHz et passe-bas 14 kHz. La bande passe-bas de 3 kHz a été choisie pour représenter une valeur prudente de la fréquence supérieure de coupure du système de télécommunication. La contribution des signaux repères monoauriculaires et spectraux a pu être observée dans la bande passe-haut de 3 kHz. Le stimulus passe-bas de 14 kHz a permis d'évaluer les écarts interauriculaires entre les signaux repères monoauriculaires et spectraux dans une combinaison qui serait disponible dans une bande plus large. En champ libre, les essais additionnels on consisté à augmenter chacune des fréquences de coupure passe-bas et passe-haut de 3 kHz à 4 kHz afin de déterminer si la fréquence des inversions en champ libre pouvait être diminuée. Sept FTAMT génériques ont été utilisées.

Selon cette première recherche, la localisation, telle que mesurée par le pourcentage moyen correct et les inversions avant/arrière, est plus précise en champ libre que dans les deux EAV. Le rendement de localisation moyen en champ libre est un peu meilleur si les fréquences de coupure passe-bas et passe-haut sont de 4 kHz plutôt que de 3 kHz. Compte tenu de ce dernier résultat, on peut aussi s'attendre à un rendement meilleur dans un EAV. Le type de FTAMT générique n'a pas influé beaucoup sur le rendement des sujets dans un EAV; la précision de la localisation avec un stimulus à large bande n'a pas été sensiblement meilleure qu'avec un stimulus passe-bas de 3 kHz. Selon ce résultat, les signaux repères spectraux sont très peu utiles pour des sources sonores situées dans le plan horizontal et une réduction de la largeur de bande du système de télécommunication à 3,5 kHz ne nuirait pas grandement à la précision de localisation de l'utilisateur dans une EAV. Le rendement de localisation a été plus faible en présence du bruit diffus ambiant d'un char Leopard que dans des conditions de calme, ce qui indique que la technologie audio 3-D ne serait pas encore très utile dans les contextes militaires très bruyants d'aujourd'hui.

Les données actuelles sont limitées au choix de positions et de stimulus dans l'espace. Il s'est avéré que l'utilisateur est plus précis dans un EAV et commet moins d'inversions de localisation avec des FTAMT personnalisées plutôt que génériques. Toutefois, les FTAMT personnalisées sont en général établies à partir de mesures biauriculaires prises dans les oreilles de l'auditeur final assis dans une chambre anéchoïde. Cela suppose un investissement majeur dans l'infrastructure et le matériel, ce qui est actuellement impossible dans la plupart des applications. Il faut poursuivre la recherche afin de choisir ou de modifier rapidement et

avec précision une FTAMT générique pour l'application visée. Il faut aussi étudier de plus près l'effet du bruit diffus ambiant sur le rendement de l'utilisateur avec des FTAMT personnalisées ou génériques. Il faut s'attaquer en particulier au problème de la largeur de la bande de télécommunication, attribuable au matériel, dans le cas de sources sonores virtuelles situées à l'extérieur du plan horizontal. Tant que nous n'aurons pas approfondi et résolu ces problèmes, il vaudrait mieux faire preuve de prudence avant d'adopter un système audio 3-D dans des applications pour des missions critiques.

Table of contents

List of figures

List of tables

Acknowledgements

The authors thank David Beevis and Sharon McFadden for helpful comments on earlier drafts of this paper. The authors also express appreciation to Sherry Van Blyderveen who served as the experimenter for all of the studies. Technical assistance was provided by Melissa Bender, Melanie Boultbee, Brian Crabtree, Phiroz Dastoor, Garry Dunn, Carolyn Michielsen, Alexandra Ostaniewicz, and Athanasia Pallas. The HRTFs used in this study were provided through the generosity of Bo Gehring, Simon Oldfield, Tim Tucker, and Fred Wightman.

Background

The successful application of three-dimensional (3-D) auditory display technology in operational military environments depends on a variety of factors. These include (but are not limited to): the faithfulness with which 3-D sound images are created, the ease of implementation, how well listeners can detect, discriminate and localize virtual auditory signals, and ultimately, the cost. Under Work Unit 6kd15, the Human-Computer Interaction, and Communications Groups at the Defence and Civil Institute of Environmental Medicine (DCIEM) have completed a variety of preliminary studies on 3-D auditory displays. This work has investigated the requirements of ear-worn transducers for realistic 3-D imaging and the comparative effectiveness of several head-related transfer functions (HRTFs). HRTFs, digital filters for synthesizing the location of a sound in virtual auditory space (VAS), are used to digitally manipulate an auditory signal. When the signal is presented over headphones, the listener experiences an illusion of spaciousness and directionality akin to free-field listening.

Several subjective attributes of diotic (the same sound presented to both ears) versus 3-D presentation have also been studied at DCIEM. These include signal detection performance in real-world masking noise and under sustained positive acceleration. Studies of speech signal discrimination with differing ear transducers, and diotic and diffuse-field maskers, have also been completed. The results of these studies suggest that the efficacy of 3-D auditory displays is at least partially maintained under adverse conditions of noise immersion and sustained +3Gz positive acceleration with respect to signal detection and discrimination.

In addition to DCIEM's commitment to continuing this research, a number of agencies within the Department of National Defence (DND) have expressed a keen interest in the air and land applications of 3-D audio display technology. The next step in the research program is to investigate sound localization in free-field and virtual auditory space (VAS) in both quiet and noisy listening environments.

To assist this investigation, DCIEM entered into a collaborative contract arrangement, #W7711-8-7455 "An Investigation into the Impact of Non-Individualized Head-Related Transfer Functions on Auditory Localization". The work performed in the contract period from May to September, 1998, is reported in this paper.

This page intentionally left blank.

Introduction

A modern combat vehicle passenger compartment or military aircraft cockpit is highly dynamic and complex. The crew often experience high workload. They must maintain situational awareness, while making quick decisions and prompt responses. A relatively new technology, the 3-D audio display, is being explored for improving crew performance in both ground and air conditions. Applications include auditory warnings (Doll, Gerth, Engelman and Folds, 1986; Calhoun, Valencia and Furness, 1987; Calhoun, Janson and Valencia, 1988), air traffic control displays (Wenzel, 1994), head-up auditory displays for traffic collision avoidance (Begault, 1993; Begault and Pittman, 1996), enhanced visual target detection and identification (Bronkhorst, Veltman and van Breda, 1996; Perrott, Cisneros, McKinley and D'Angelo, 1996; D'Angelo, Bolia, McKinley and Perrott, 1997), and speech intelligibility (Begault and Erbe, 1994; Ricard and Meirs, 1994; Ericson and McKinley, 1997). It has been proposed that a 3-D auditory display can support situational awareness and spatial orientation by providing veridical spatial cues to the positions of targets, threats, and beacons (Doll et al., 1986; Furness, 1986; Stinnett, 1989). Arrabito, Cheung, Crabtree and McFadden (2000) found that the detection level of a pulsed signal while subjects were under sustained positive G-stress was significantly lower in VAS compared to a diotic presentation. During an in-flight study, pilots reported that a 3-D audio display decreased target acquisition time and visual workload while increasing communication capability and situational awareness (McKinley and Ericson, 1997).

The effectiveness of a 3-D audio display depends on the listener's ability to discriminate and localize various sources of information in auditory space. Spatialization of an auditory signal over headphones is accomplished by digitally filtering the signal with head-related transfer functions (HRTFs). These HRTFs encode the binaural and spectral cues used in sound discrimination and localization. It has been argued that a listener's ability to localize virtual sound is more accurate when using HRTFs measured from his/her own head (personal) compared to HRTFs measured from a different head (generic) (Wightman and Kistler, 1989b; Wenzel, Arruda, Kistler and Wightman, 1993; Carlile and Pralong, 1994). The investigators of these studies have shown that generic HRTFs contribute significantly to reversals (i.e., perceiving the mirror image of the presented sound source). An example of a front-back reversal occurs when a listener locates a sound position in the rear hemifield at 135° azimuth when the actual sound source was presented in the front hemifield at 45° azimuth. Back-front reversals also occur but are less frequent than front-back (Oldfield and Parker, 1984a, b, 1986). Front/back[1] reversals are believed to result from the inherent ambiguity in the interaural time of arrival (ITD) and interaural level difference (ILD) cues. A given interaural difference specifies a number of positions in space. If the head is kept stationary, then a given ITD will not be sufficient to define uniquely the position of the sound source in space. There is a cone-of-confusion (Mills, 1972) such that any sound source on the surface of this cone would give rise to the same ITD. The same is also true for ILD. For example, a cone-of-confusion for a particular interaural time delay is illustrated in Figure 1.

The frequency of reversals in free-field and in VAS has been calculated in laboratory settings (Oldfield and Parker, 1984a, b, 1986; Wightman and Kistler, 1989b; Wenzel et al., 1993). Investigators have shown that the frequency of front/back reversals is subject dependent. When using generic HRTFs, some listeners have a front/back reversal rate as high as 50% in VAS compared to 43% in the free-field while other listeners have a front/back reversal rate as low as 10% in VAS compared to 2% in free-field (Wenzel et al., 1993). There is a smaller variability of front/back reversals in VAS when using personal HRTFs (Wightman and Kistler, 1989b). The authors of this paper are not aware of published studies that report on the frequency of reversals occurring in day-to-day life. In the event that the results of Wenzel et al. (1993) are representative of real-world listening experiences, a front/back reversal rate as high as 50% in VAS is clearly unacceptable.

[1] In this paper, the term "front/back" denotes both front-back and back-front reversals.

For example, there is no tolerance for reversals if virtual sources are to cue armoured vehicle or air crew to the spatial location of a potential lethal threat.

Front/back reversals are largely resolved by the presence of spectral cues. Spectral cues are contained in the frequency region above 4 kHz and are encoded by the head, pinnae and upper torso (Blauert, 1983). The spectral shaping by the pinnae is highly directional dependent (Shaw, 1974, 1975). In VAS, pinnae cues are also responsible for the externalization of the acoustic image outside of the listener's head (Plenge, 1974; Durlach, Rigopulos, Pang, Woods, Kulkarni, Colburn and Wenzel, 1992). The absence of spectral cues degrades accuracy in binaural localization (Ivarsson, de Ribaupierre and de Ribaupierre, 1980; Musicant and Butler, 1984a, b, 1985; Butler, 1986; Middlebrooks, 1992, 1999b; Bronkhorst, 1995), monaural localization (Belendiuk and Butler, 1975; Butler and Planert, 1976; Bloom, 1977a, b; Butler and Flannery, 1980; Ivarsson et al., 1980; Flannery and Butler, 1981; Musicant and Butler, 1984b, 1985; Butler, 1986; Butler, 1987), localization on the vertical plane (Bloom, 1977a; Watkins, 1978; Middlebrooks, 1999b), and localization on the median sagittal plane (Hebrank and Wright, 1974; Butler and Planert, 1976; Bloom, 1977b; Butler and Belendiuk, 1977).

The role of spectral cues and type of HRTF are two factors that could influence localization accuracy in real-world applications incorporating virtual auditory cueing via the communication system. The bandwidth of the communication system of tracked combat vehicles and aircraft has a typical upper cutoff frequency between 3.5 and 4 kHz (Patterson, 1982; Ericson and McKinley, 1997; King and Oldfield, 1997; Nixon, Anderson, Morris, McCavitt, McKinley, Yeager and McDaniel, 1998). If the primary cues afforded by the bandwidth of the communication system are the differences in the time of arrival and level at the two ears then one could safely assume a greater incidence of front/back reversals. If virtual sources are to be used in a general-purpose 3-D audio display under critical conditions, then the HRTFs should be optimized for the targeted application. Greater localization accuracy would be achieved if subjects used personal HRTFs (Wightman and Kistler, 1989b; Wenzel et al., 1993; Carlile and Pralong, 1994). However, it is not presently practical or affordable to measure HRTFs for each potential listener. The attributes of signal bandwidth and generic HRTFs must be studied in order to determine the listener's ability to accurately localize the virtual audio signal.

Study goal

The goal of the present study was to investigate the impact of generic head-related transfer functions and signal bandwidth on auditory localization in the horizontal plane in the quiet and in the presence of operational noise. Seven generic HRTFs were used for the spatialization of the acoustic stimulus for presentation in virtual auditory space. The acoustic stimulus was white noise band-limited in one of three ways: low-pass 3 kHz, high-pass 3 kHz, and low-pass 14 kHz. The low-pass 3 kHz was chosen to reflect the conservative upper cutoff frequency of the communication system. The contribution of monaural and spectral cues could be observed in the high-pass 3 kHz. The low-pass 14 kHz stimulus allowed an assessment of the interaural difference, monaural, and spectral cues in a combination that would be available in a broader bandwidth. Testing was performed in VAS and in free-field. The VAS and free-field speaker array configuration consisted of eight speaker positions, spanning 360° around the listener. Four experiments were conducted to meet the goals of the study:

1. The comparison of localization accuracy in VAS in quiet (71 dB(A)) and in the presence of diffuse ambient Leopard tank noise (approximately 110 dB(A)). These conditions are described in Experiments 1 and 2, respectively.

2. Testing in the free-field which partially served to psychoacoustically validate the VAS conditions. The free-field testing is reported in Experiments 3 and 4.

3. Replication in the free-field of the acoustic stimuli conditions used in VAS (Experiments 1 and 2). This is described in experiment 3.

4. In order to determine if the frequency of reversals in free-field could be decreased, the low- and high-pass 3 kHz cutoff frequencies were changed to low- and high-pass 4 kHz, respectively. This is described in Experiment 4.

VAS localization in quiet and in noise

Experiment 1: VAS localization in quiet

Method

Subjects

Two male and three female subjects voluntarily participated in this study. The subjects ranged in age from 21 to 53 years, with a mean age of 33. A Békésy audiometric test was administered to each subject. All participants had less than a 20 dB bilateral hearing loss at frequencies between 125 Hz and 8 kHz, and reported no history of hearing abnormalities. Four of the subjects were in-house employees. The fifth was recruited from the general population outside of DCIEM. With the exception of the first author (RA), who participated as a subject, none of the subjects had previously participated in psychoacoustic studies and were naive regarding the purpose of the experiment. The DCIEM Human Ethics Committee approved the experimental protocol and informed consent was obtained from the subjects; subjects were given stress allowance in accordance to guidelines established by DND and DCIEM.[2]

Stimuli

The stimulus was 300 ms of white noise with a 50 ms linear onset/decay, band-limited in one of three ways: low-pass 3 kHz, high-pass 3 kHz and low-pass 14 kHz. These were chosen to allow an assessment of the effectiveness of binaural and spectral cues.

Head-related transfer functions

The transfer functions of the head and external ears can be captured by presenting an impulsive broadband sound at various locations in the vicinity of the head in a free-field. Recordings of the sound source are measured by placing a microphone in the ears of an acoustic mannequin (Plenge, 1974; Doll et al., 1986) or the ear canals of a human (Butler and Belendiuk, 1977; Wightman and Kistler, 1989a; Bronkhorst, 1995). Transfer functions measured in this way have come to be known as head-related transfer functions (HRTFs). An HRTF includes the effects of diffraction by the head, neck, and upper torso, in addition to spectral shaping by the pinnae. The binaural impulses are then implemented into a pair of digital filters for use in a 3-D audio system using convolution techniques (Oppenheim and Schafer, 1989). When HRTFs are applied to an arbitrary signal and presented to the listener over headphones, a virtual target that appears to originate from the location of the original sound source is heard (Wightman and Kistler, 1989b; Carlile and Pralong, 1994; Bronkhorst, 1995; Moller, Sorenson, Jensen and Hammershoi, 1996). A set of transfer functions can be compiled for a wide range of sound-source locations, synthesizing a virtual auditory environment. For a layman's description of this procedure, the interested reader is

[2] This statement also applies for the subjects who participated in Experiments 2, 3, and 4.

referred to Leong, Tucker and Carlile (1996). A more advanced treatment of digital signal processing techniques can be found in Oppenheim and Schafer (1989).

Seven different generic HRTFs were used in this study to spatialize the acoustic stimulus in virtual auditory space. They will be denoted in this paper as "A", "F", "K", "R", "S", "T", and "W". The measurement techniques and psychoacoustic validation for HRTF "A" are reported in Pralong and Carlile (1994) and Carlile and Pralong, (1994), respectively. The measurement techniques and psychoacoustic validation for HRTF "R", "S" and "W" are reported in Wightman and Kistler (1989a, b). The measurement techniques and psychoacoustic validation for HRTF "F" and "T" were not available. These HRTFs were provided to the present authors by Bo Gehring of Focal Point 3-D Audio and Tim Tucker of Tucker-Davis Technologies, respectively. Unlike the aforementioned HRTFs that were measured on humans, HRTF "K" was measured on the Knowles Electronic Mannequin for Acoustic Research (KEMAR), described by Burkhard and Sachs (1975). The corresponding measurement technique is reported in Gardner and Martin (1994). With the exception of HRTF "R" and "S", which were measured on the first author[3], all other HRTFs were measured on individuals who did not participate in this study. All HRTFs were implemented in a Tucker-Davis Technologies system, which was used in the course of this study (see below).

Apparatus and calibration

Subjects were tested individually while seated on a chair in an Industrial Acoustics Company (IAC) double-wall sound attenuation booth located at DCIEM. The ambient level for all frequencies in the booth was less than the maximum allowed for open-ear headphone testing (ANSI-S3.1, 1991). The booth contained a window and an intercom, which allowed the experimenter to monitor the subject.

Sound localization was assessed via customized software in conjunction with Tucker-Davis Technologies (TDT) equipment. The TDT was used for presenting the acoustic stimulus in VAS and for collecting responses from the subject. The TDT system consists of a suite of digital and analog audiometric equipment, controlled by a 486 personal computer (PC) that served as the host computer in this study.

A block diagram of the apparatus is shown in Figure 2. The output of a Brüel and Kjær (B&K) Type 1405 noise generator was routed to the TDT PD1 that filtered and spatialized the acoustic stimulus. Signals were generated with 16-bit precision at a sampling rate specified by the HRTF ("A": 40 kHz; "K": 44.1 kHz; "F", "R", "S", "T", and "W": 50 kHz). The output from the TDT PD1 (powerdac) was routed to a TDT FT6 two-channel anti-aliasing low-pass 30 kHz filter. This removed the alias products which result from digital-to-analog conversion. The left- and right-channel filtered outputs were routed to two TDT PA4 programmable attenuators used to set the amplitude of the acoustic stimulus. The outputs from the TDT PA4s were routed to a TDT SW2 (cosine switch) that controlled the onset and decay of the stimulus. The outputs from the TDT SW2 drove the Stax SRM-T1 headphone amplifier. The stimulus was then presented over the Stax electrostatic headphones (model SR-Λ Signature). A custom-made response box for subjects to make

[3] As the first author participated as a subject in the present study, HRTFs "R" and "S" allowed an assessment of sound localization using personal HRTFs. It was not possible to measure HRTFs for any of the other subjects who participated in this study.

localization judgements and a set of light emitting diodes (LEDs) used to cue subjects, were connected to the TDT PI2 (parallel interface).

The calibration of the acoustic stimulus was performed with a Brüel and Kjær (B&K) 4134 1/2 inch pressure microphone mounted within a shock-mounted flat-plate coupler. The Stax earcup was pressed against the coupler plate. The microphone recorded the signal output from the Stax earcup. A 16 second sample of each of the conditions (HRTF x acoustic stimulus) spatialized at 0° azimuth comprised the signal output. The signal from the microphone was fed to a B&K 2133 frequency analyzer. The left and right outputs from the Stax earcups were separately measured and then averaged. The TDT PA4 was used to set the sound level at 71 dB(A). The deviation from the 71 dB(A) was ± 0.3dB(A). All other azimuth positions for each of the HRTF x acoustic stimulus conditions were then attenuated by the same value.

Experimental design

A 7 (HRTF) x 3 (acoustic stimulus) x 8 (azimuth) x 4 (session) within-subject repeated measures design was employed to assess a subject's ability to localize the acoustic stimulus in virtual acoustic space. The acoustic stimulus was spatialized at one of eight static azimuth positions on the horizontal plane at 45° intervals starting at 0° azimuth. In this paper, azimuth increases clockwise on the horizontal plane, with 0° positioned directly in front of the listener.

A block was comprised of one of the three acoustic stimuli and one of seven HRTFs. A block consisted of 8 practice trials followed by 40 experimental trials. Each azimuth position was presented once in the practice trials and five times in the experimental trials. A session contained 21 blocks (7 HRTFs x 3 stimulus conditions). A Latin Square design was used to counterbalance azimuth position and blocks across subjects and sessions.

Procedure

Subjects were individually tested in the IAC listening booth. Each subject was given a dummy training block before starting the experiment. Subsequently, data collection commenced. Each experimental trial began by flashing a 500 ms green LED on the cue box located on the wall of the IAC booth and positioned at the subject's eye level. This was followed by a 500 ms delay prior to the presentation of the acoustic stimulus. The subject's task was to indicate the perceived location of the acoustic stimulus. Subjects made localization judgements by pressing a button on the response box that was situated on their lap. The buttons were arranged in the same configuration as the virtual speaker array. An 8-alternative forced-choice paradigm in which the response alternatives corresponded to the source alternatives was used. This method minimizes response bias (Green and Swets, 1966). At the off-set of the acoustic stimulus, the subject was given a maximum of ten seconds to make a localization judgement. If a response was not made during this time period, the trial was scored as a miss. Subsequently, the next trial was presented. The azimuth was recorded. No feedback was given to the subjects regarding the accuracy of their localization judgements. Following the completion of a block of trials, subjects proceeded onto the next block until the 21 blocks were completed. Five minute breaks were given at approximately 30 minute intervals. The duration of each session was approximately 90 minutes. Subjects completed four sessions, each on a different day. The subjects were monitored by the experimenter through a window in the sound listening

booth, in addition to an intercom located in the booth. Subjects were debriefed following the completion of the study.

Experiment 2: VAS localization in noise

Method

Subjects

Two male and three female subjects voluntarily participated in this study. The subjects ranged in age from 17 to 35 years with a mean age of 23. A Békésy audiometric test was administered to each subject. All participants had less than a 20 dB bilateral hearing loss at frequencies between 125 Hz and 8 kHz, and reported no history of hearing abnormalities. Three of the subjects (KA, RA and SV) were in-house employees who also participated in Experiment 1; the others were recruited from the general population outside of DCIEM.

Stimulus and masker

The stimulus was the same as in Experiment 1. Diffuse ambient noise (approximately 110 dB(A)), simulating a Leopard tank traveling at approximately 30 km/h on hard standing, was used to mask the stimulus. The spectrum of the masker is shown in Figure 4.

Apparatus

The same apparatus as in Experiment 1 was used with the exception of the Stax headset. The Stax headset was replaced with a Racal Armored Vehicle Headset (AVH), manufactured by Racal Acoustics Limited (Middlesex, England). The AVH headset is used by military personnel to protect their hearing and to communicate. The AVH headset incorporates active noise reduction (ANR), a technique for electronically reducing noise levels at the ears of the observer by means of interfering sound waves. The net result is a partial cancellation of noise at frequencies up to approximately 1000 Hz. ANR headsets are used in fixed- and rotary-wing aircraft and tracked armored fighting vehicles. In these environments, the level of ambient noise is typically in excess of 100 dB SPL; this level of noise degrades aural communication and is potentially hazardous to hearing. Although noise-excluding earcups are an integral part of most headgear (flight and vehicle helmets and headsets), their passive attenuation at low frequencies is limited. The frequency response and amount of attenuation of the Racal AVH are illustrated in Figures 5a and 5b, respectively.

Subjects were individually tested while seated on a chair in the DCIEM Noise Simulation Facility. The Noise Simulation Facility is a large reverberant chamber (11 X 6 X 3 m) in which the sound produced by equipment such as helicopters and tracked vehicles may be faithfully reproduced. An array of loudspeakers placed at one end of the chamber produces the desired sound field, at levels approaching 130 dB SPL over a bandwidth of 15 to 20,000 Hz, except at very high frequencies. The facility also meets the ANSI S3.1-1991 standard with respect to sound field uniformity and diffusivity for open-ear testing. The low-frequency sound is reproduced by 16 18-inch Gayne Electronics loudspeakers, which are housed in 8 closed boxes and by 4 Servo-Drive Bass Tech 7 front-loaded horns. Mid and

high frequency sound is produced by four Electro-Voice Deltamax systems employing reflex-loaded direct radiator speakers and horn-loaded compression drivers. The subject was seated facing the speakers and was positioned 5.3m from the center loudspeaker. An easily accessible "kill" switch was attached to the underside of the subject's chair so that the noise could be immediately turned off if the subject so desired. The experimenter monitored the subject via a video camera from within the control room of the Noise Simulation facility.

Experimental design and procedure

The experimental design and procedure were similar to Experiment 1 with the following exceptions. The sound level of each of the acoustic stimulus conditions was increased to 75 dB(A), as measured at the 0° azimuth position, to allow for the acoustic stimulus to be comfortably heard in the presence of the diffuse masker. Prior to beginning the experiment and after each rest period, the experimenter saturated the Noise Simulation Facility with the Leopard tank noise for approximately 90 seconds to allow the subject to be acclimatized to the noise level. Subsequently, testing began. Subjects took five minute breaks approximately every 20 minutes. During the rest period, the Leopard tank noise was off.

Results

Before proceeding, it is necessary to discuss a point that merits attention from a safety perspective given that directional cueing may be used in critical mission applications. This concerns the occurrence and treatment of localization judgements that result in reversals. In localization studies, reversals are commonly resolved by coding the subject's response as if it were indicated in the correct hemisphere (Oldfield and Parker, 1984a; Butler, 1986; Wightman and Kistler, 1989b; Begault and Wenzel, 1993; Wenzel et al., 1993). Clearly, resolving reversals in this manner for critical mission applications could be fatal. There is no tolerance for reversals if virtual sound sources are to cue the crew to the spatial location of a potential lethal threat, such as another armored vehicle and/or aircraft. However, for the purpose of comparison with other published reports, the data were partitioned into a) corrects, b) adjusted corrects for reversals (i.e., front-back, back-front, left-right, right-left, and diagonal), and c) errors (i.e., localization judgements that are neither correct or reversed). Table 1a shows the distribution of localization judgements, classified as corrects, adjusted corrects, errors, and front-back and back-front reversals, as a function of stimulus type and HRTF. Table 1b shows these responses for the noise condition. Because the occurrence of left-right, right-left and diagonal reversals was extremely small (1.1% and 3.9% for Experiments 1 and 2, respectively, when averaged across stimulus and HRTF) compared to the total number of trials in each experiment (16,800), they have been omitted from the tables but they were included in the calculation of the adjusted corrects. In order to compare these results with other published reports, the "adjusted" correct performance (whereby a reversal is reclassified as a "correct" localization judgement) is included. In general, the greatest number of reversals (collapsed across stimuli and HRTFs) occurred for the front-back spatial positions. There were more reversals under the quiet condition than the noise condition across HRTFs and stimulus bandwidth. In contrast, there were more errors under the noise than the quiet condition.

In the quiet condition, the ability to accurately localize a sound source was affected by stimulus ($F(2,8) = 10.08$, $p < 0.01$) and HRTF type ($F(6,24) = 3.65$, $p < 0.01$). A Scheffe post hoc analysis (alpha = 0.05) failed to reveal any significant effect due to HRTF. However, a Scheffe post hoc analysis (alpha = 0.05) revealed that localization performance using the broadband (BB) or low-pass (LP) stimuli was significantly better than the high-pass (HP) stimulus. In addition, while there was no significant effect due to azimuth, there was a propensity for subjects to localize more accurately when the sound was presented between 90° and 225° azimuth.

Unlike localization under the quiet condition, subject performance was not significantly affected by HRTF or stimulus when localizing under the noise condition. However, it was affected by azimuth ($F(7,28) = 3.68$, $p < 0.01$). In addition, there was a significant effect due to session ($F(3,12) = 6.71$, $p < 0.01$). A Scheffe post hoc analysis (alpha = 0.05) showed a significant effect for sessions only. This indicates that subject performance improved with practice over sessions.

There was a significant effect of sound localization accuracy between quiet and noise conditions when collapsing across HRTFs and signal bandwidth ($F(1,8) = 16.33$, $p < 0.01$). A Scheffe post hoc analysis at the 0.05 alpha level revealed that localization performance was significantly poorer in noise (42.9%) than in quiet (46.4%). Three subjects (KS, RA and SV) participated in both the quiet and noise conditions. Localization performance under quiet and noise conditions for these subjects was similar, especially between 90° and 225° azimuth. However, at the azimuth locations of 0°, 45°, 270° and 315°, performance under the noise condition was more degraded then under the quiet condition. It should be noted that "subject" was treated as a between-subject factor in spite of subjects KS, RA and SV who participated in both quiet and noise conditions. Separate analysis on the within- and between-subject factors could have been performed but the number of participants in each group was too small thus leading to interpretations which would have very little practical significance.

The issue of interaction deserves special mention. Some of the interactions between factors (testing environment, HRTFs, stimulus bandwidth, azimuth position and session) were statistically significant ($p < 0.05$). However, the Scheffe multiple comparison test (which makes a conservative adjustment for multiple testing) at the 0.05 alpha level failed in many instances to identify which group mean(s) differed from the others. In addition, they were not large enough or consistent enough to be of any practical significance. For this reason, the treatment of interactions for the VAS experiments and those for the following two free-field experiments described in the present paper are omitted from subsequent discussions.

One feature of sound localization in VAS observed in the present study was the degree of localization performance as a function of HRTFs. Figure 6 illustrates HRTF differences, observed under the quiet stimulus condition as a function of azimuth location for LP (Figure 6a), HP (Figure 6b), and BB (Figure 6c) stimulus conditions. Similarly, Figure 7 shows HRTF differences in VAS under the noise condition for LP (Figure 7a), HP (Figure 7b), and BB (Figure 7c).

In general, localization accuracy for the LP stimulus was better under the VAS-quiet condition than the VAS-noise condition for all HRTFs. Performance with HRTF "T" was the same while the largest performance difference occurred with HRTF "S" (14.2%). With the exception of HRTF "F", localization performance with the HP stimulus was better under the VAS-noise versus the VAS-quiet condition. The smallest difference in performance was observed with HRTF "S" (0.8%), while the largest was observed with HRTF "T" (7.1%). For the BB stimulus, performance for all HRTFs was better under the VAS-quiet as compared to the VAS-noise condition. The smallest difference in performance between the VAS-quiet and VAS-noise condition was observed with HRTF "T" (0.9%), while the largest occurred with HRTF "W" (17.3%).

There was a significant effect of reversals between quiet and noise conditions when collapsing across HRTFs and signal bandwidth ($F(1,8) = 27.95$, $p < 0.01$). A Scheffe post hoc analysis at the 0.05 alpha level revealed that there were significantly fewer reversals in noise (20.4%) than in quiet (24.6%). In keeping with the common practice of coding the subject's response as if it were made in the correct hemisphere (Wightman and Kistler, 1989b; Wenzel et al., 1993), Figure 8 shows the HRTF that produced the best and worst performance for correct and adjusted correct performance under the quiet condition. For localization of LP (Figure 8a) and BB (Figure 8c) stimuli, the best correct performance occurred with HRTF "S". For the HP stimulus (Figure 8b) the best correct performance occurred with HRTF "R". For most stimulus conditions, the worst performance for correct responses occurred when subjects were localizing with HRTF

"A". For best adjusted correct responses, there was a tendency for HRTF "F" to elicit the most accurate performance for LP and HP stimulus conditions. The worst adjusted correct performance occurred for HRTF "R" with the HP and BB stimuli, and for HRTF "T" with the LP stimulus. As was observed with correct performance (Table 1), adjusted correct performance was significantly affected by stimulus ($F(2,8) = 18.61$, $p < 0.01$). A Scheffe post hoc analysis (alpha = 0.05) revealed that, as with correct performance, adjusted correct performance using a HP stimulus was significantly less accurate than when localizing either LP or BB stimuli. There was no significant difference in performance with LP or BB stimuli. There was little difference between the correct and adjusted correct performance for LP as compared to HP and BB stimuli with the HRTFs that produced the worst performance. In addition, unlike correct performance, there was a significant effect due to azimuth position ($F(7,28) = 3.68$, $p < 0.01$). However a Scheffe post hoc analysis at the 0.05 alpha level did not reveal any significant difference between the azimuth locations.

The best and worst HRTFs for both correct and adjusted correct performance under the noise condition are shown in Figure 9a for the LP stimulus, Figure 9b for the HP stimulus, and Figure 9c for the BB stimulus. On average, HRTF "T" yielded the most accurate performance for both correct and adjusted correct responses across all three stimulus conditions. Both HRTF "A" (correct) and HRTF "R" (adjusted correct) produced the worst performance for the HP and BB stimulus conditions. Adjusted correct performance was significantly affected by HRTFs ($F(6,24) = 3.00$, $p < 0.02$). This contrasts with the correct performance described above for the VAS-quiet condition, where performance was affected by sessions only (i.e., there was a practice effect).

There were also differences in localization performance across subjects. For example, Figure 10 shows subjects' performance in the LP stimulus under the quiet condition for correct (Figure 10a) and adjusted correct (Figure 10b) performance. Greater differences were observed between subjects for correct as compared to adjusted correct performance. Figure 11 illustrates differences in subjects' performance under noise conditions for the HP (Figure 11a) and LP (Figure 11b) stimuli. As under the quiet stimulus condition, greater differences were observed between subjects in the correct (Figure 11a) as compared to the adjusted correct (Figure 11b) performance.

The localization judgements that result in errors merit some attention. Under the VAS-quiet stimulus condition, there was a tendency for more errors to occur with HP (32.7%) than for either LP (28.5%) or BB (25.6%) stimuli. Performance under the noise condition did not yield much difference across stimulus conditions. There was a significant effect of errors between the quiet and noise conditions when collapsing across HRTFs and signal bandwidth ($F(1,8) = 103.08$, $p < 0.01$). A Scheffe post hoc analysis at the 0.05 alpha level revealed that there were significantly fewer localization errors in quiet (28.9%) than in noise (36.4%).

One final aspect of sound localization in VAS warrants some attention; performance using one's "personal" (i.e., own) HRTF compared with using generic HRTFs. In the present study, only subject RA was tested with his own HRTF "R" and "S", along with the five other generic HRTFs. When collapsing performance across azimuth, his performance was slightly worse compared to the other subjects in the quiet condition and slightly better than that of others in the noise condition (e.g., performance is shown for the BB stimulus for quiet and noise in Figure 12a and b, respectively). In the VAS-quiet condition using HRTF "S" (measurement made at the open ear canal), subject AK had a higher average of corrects compared to subject RA: 65.6%, 63.1% and 88.1% in LP, HP and BB, respectively, versus 65.6%, 52.5% and 78.1%, respectively. Similar results were also obtained with HRTF "R" (measurement made at the blocked ear canal). On the other hand, in the VAS-noise condition subject RA obtained the highest average of corrects using HRTF "R" and "S".

Discussion

Signal bandwidth

Different outcomes were observed in the quiet and noise conditions as a function of signal bandwidth. In the quiet condition, localization accuracy in the LP stimulus was not significantly poorer than the BB stimulus condition. Localization performance in the noise condition was not significantly affected by signal bandwidth. Caution is required in concluding that these results suggest that localization judgements in VAS on the horizontal plane may not be significantly affected by the restricted bandwidth of the communication system of present-day armored vehicles and aircraft which have a typical upper cutoff frequency between 3.5 and 4 kHz (Patterson, 1982; Ericson and McKinley, 1997; King and Oldfield, 1997; Nixon et al., 1998). The caution stems from two primary factors. Not all subjects participated in both VAS conditions, and two different headsets were used. In the latter case, the Stax headset used in the VAS-quiet condition is not a hearing protection device (HPD) and thus could not have been utilized in the presence of the diffuse ambient Leopard tank noise.

In general, there were more front-back reversals compared to back-front reversals, as illustrated in Tables 1a (VAS-quiet) and b (VAS-noise) regardless of the signal bandwidth. It is common for more front-back than back-front reversals under both VAS (Wightman and Kistler, 1989b; Wenzel et al., 1993) and free-field (Oldfield and Parker, 1984a, b, 1986) conditions. On average, subjects made more front/back reversals in the HP condition than in either the LP or BB conditions. The occurrence of left-right, right-left and diagonal reversals were minimal, in agreement with Bronkhorst (1995).

The poorer localization accuracy in the VAS-quiet condition, on average, for the HP stimulus condition agrees with the observation that the sole presence of spectral cues are not a sufficient condition for making a correct front/rear judgement (Asano, Suzuki and Sone, 1990). For example, Asano et al. (1990) investigated the spectral cues in the external ear transfer functions that aid in median plane localization. Testing was performed in VAS using personal HRTFs. They found that in order to achieve accurate front/back localization judgements, subjects required the presence of frequencies below 2 kHz and that the simplification of spectral cues in high frequencies had little influence on front/rear judgement. In particular, the authors conclude that: (1) information from macroscopic patterns in the high frequency region is necessary for front/rear judgement, while microscopic spectral patterns in the high frequencies convey little information for front/rear judgement; and (2) information from microscopic spectral cues below 2 kHz is not sufficient but necessary for correct front/rear judgement, as long as the signal is broadband and contains components below 2 kHz. The fact that the LP stimulus was not significantly different from the BB stimulus in the present study may be explained by noting that localization judgements were tested only on the horizontal plane. The presence of interaural cues are robust for localization judgements on the horizontal plane (Blauert, 1983). In the study by Asano et al. (1990) testing was performed off the horizontal plane and thus the absence of spectral cues degrades vertical plane localization (Bloom, 1977; Watkins, 1978; Middlebrooks, 1999b).

Quiet versus noise

Subjects made fewer errors and more reversals in the quiet condition (Experiment 1), as compared to the noise condition (Experiment 2). When averaging across HRTFs and stimuli for the quiet condition, the number of errors and reversals was 28.9% and 26.7%, respectively. In the noise condition, these values were 36.4% and 20.7%, respectively. The poorer performance in noise

compared to the quiet condition is primarily attributed to the presence of the diffuse ambient Leopard tank noise, presented at approximately 110 dB(A). This was chosen to reflect the noise levels observed in the Leopard tank which range from 98-112 dB(A) as measured by Forshaw and Crabtree (1983). Although the present investigators did not measure the subject at-ear signal-to-noise ratio (SNR) of the acoustic stimuli, all subjects reported that the signal was clearly audible independent of HRTF, acoustic stimulus and azimuth position.

The presence of more errors in the noise condition might also be explained by noting that localization on the horizontal plane primarily depends on interaural differences in time and level (Blauert, 1983). With respect to the interaural level difference cue, judgement of direction tends to be toward the side of the listener's head receiving the louder stimulus. This cue is particularly relevant for studies of the effects of HPDs on sound localization. The Racal AVH is designed to reduce the at-ear sound level, and thus equal attenuation for both ears would not change the interaural level difference. However, variations in the effectiveness of sealing (due to poor fit, deterioration or dislodging) might produce changes in the pattern of level differences and make this cue unreliable. A poor seal on the Racal headset will also cause the ANR mechanism to "howl" and thereby produce additional unwanted noise at the listener's ear. Together these would lower the SNR at the ear with the breached seal and thus distort the spatial position of the presented sound source. The subject would perceive the direction of the stimulus from the side with the louder cue (i.e., the ear with the non-breached seal). This phenomena coupled with the sound level of the Leopard tank noise could explain the lack of a significant effect of HRTF and acoustic stimulus.

These observations are in partial agreement with the findings of Good, Gilkey and Ball (1997). They found that localization in the free-field can be severely degraded by the presence of a single masking sound. The impact of the degradation depends on both the SNR and the location of the masker. These investigators also suggest that the detectability of a signal may serve as a localization cue. Subjects could limit the set of potential locations from which the signal could have been presented. For example, in the case of a single masker, when the masker is in front of the listener and a highly detectable signal is heard, it is likely that the listener will perceive the direction from the left, right or above than from in front or behind. Such cues would not, in general, be available to listeners in a real-world environment. More specifically, in the present study, the signal was presented at one of eight possible locations on the horizontal plane. In a real-world situation, listeners would not typically know the loudness or direction of the sound source and thus would not have a basis for determining whether its detectability is different than "normal". Moreover the noise level and spectrum will vary in the crew compartment of a fighting vehicle at different seating positions (Forshaw and Crabtree, 1983). Thus the benefits of a HPD may not be the same amongst crew members, possibly leading to further at-ear noise levels which may distort spatial cues.

The only published report known to the investigators concerning the testing of uncorrelated ambient noise on user performance in VAS is reported by McKinley and Ericson (1997). These investigators measured the minimum audible angle (MAA). The MAA is the calculation of the minimally discriminable separation of two sound source positions (Mills, 1972). The MAA was measured on the horizontal plane. Testing was performed in the free-field and in VAS in the presence of 115 dB SPL ambient pink noise, which was used to replicate the acoustic levels in high performance aircraft cockpits. In the latter condition, the binaural Bose PRU-57 military active noise reduction headset was used. The KEMAR HRTFs were employed for the spatialization of the stimulus. The investigators found that the MAA was 4-5° in the free-field condition versus 6-7° in the noise condition. Extrapolations from the above study to the present study are difficult due to variations in methodology and test parameters.

Subject variability

Several factors are known to affect the spatial synthesis of HRTFs, which could account for the variability in subject performance as illustrated in Figures 9 and 10. For example, the measurement techniques of HRTFs differ across laboratories and are motivated by the different goals of the investigators (see references 2, 3, and 5-27 of Moller, Sorensen, Hammershoi and Jensen, 1995). Some of the parameters that vary significantly in the measurement of HRTFs are the type of test stimulus (e.g., sinusoidal tones or noise bursts), the point in relation to the ear canal where the measurement is made (e.g., at the blocked ear canal or a point somewhere along the ear canal), and the number of source positions.

Variability of HRTFs across individuals using the same HRTF measurement technique has been reported by other investigators (Shaw, 1965; Wightman and Kistler, 1989a). An example of HRTF variability is illustrated in Figure 13. The HRTFs were measured from one ear of two people in the laboratory of F.L. Wightman using techniques described in Wightman and Kistler (1989a). The source direction was kept constant at 90° azimuth while elevations ranged every 20° from -60° to +80°. As observed, there is substantial intersubject variability in the HRTF for a single source position. This is expected given differences in head size and pinnae shape. In the present study, intersubject variability for a given HRTF is illustrated in Figures 10-12. Some factors that could account for this variability are discussed below.

Generic head-related transfer functions

Prior to discussing localization performance of the HRTFs used in this study, it is first necessary to mention some limitations imposed on the investigators. Given the seven generic HRTFs used in the course of this study, it should be mentioned that there were no elements in the signal paths to "colour" or modify the signals. The HRTFs were treated as "black boxes" in the sense that the parameters were not altered in any way. User performance is therefore strictly based on the characteristics of the HRTFs. The present investigators cannot report on the characteristics of these HRTFs that could have contributed to varying perceptions among the subjects. This stems primarily from the absence of information on the subjects used for the measurement of the HRTFs and on the detailed procedures of their measurements. If such information would have been available then the present authors could have measured the subjects' pinnae and bi-tragion distance. This data could have been compared with the corresponding data of the listeners for whom the HRTFs were measured and thus used to partially account for subjects' localization performance.[4]

The present authors acknowledge that the spatial synthesis process of virtual sound sources is not solely dependent on the selection of HRTFs (either generic or personal). The headphones contribute to the total transmission and are often equalized to better reproduce the binaural synthesis process (Wightman and Kistler, 1989b; Moller, 1992; Carlile and Pralong, 1994; Bronkhorst, 1995). Headphone equalization is a digital filtering procedure used to cancel distortion caused by headphones and resonance effects of the listener's ears. However, headphone equalization is specific to the headset and the end-listener. The procedure for equalizing the headphones is similar to that for the measurement of HRTFs. This in turn makes the measurement procedure impractical in many applications. It should also be noted that headphone placement on the listener's head is rarely the same and, hence, the headphone equalization step is not always exact. It may then be argued that the headphone equalization step may be more detrimental than useful. For these

[4] The effect of mismatched pinnae and head size on localization accuracy will be discussed later on in this paper.

reasons, and given the near flat frequency response of the Stax headset used in this study (see Figure 3), the present investigators did not equalize the headphones.

In the present study it is not possible to make a general statement regarding localization accuracy based on HRTF type (generic and personal HRTFs). This is due to only one subject (RA) who used his own HRTFs in Experiments 1 and 2. A comparison could have been made if all subjects were tested with their own HRTFs in addition to generic HRTFs. It was not technically feasible to measure the HRTFs of all subjects as DCIEM does not have the required facilities and equipment to carry out such measurements.[5] Nevertheless, it is worth noting subject RA's performance compared to the other subjects. The localization performance of subject RA using HRTF "R" and "S" did not always have the highest localization accuracy for all azimuth positions. It is not uncommon for listeners not to perform best with his/her own HRTFs (Butler and Belendiuk, 1977). On the other hand, in the VAS-noise condition subject RA obtained the highest average of corrects using HRTF "R" and "S". In this instance, subject RA's higher localization accuracy might be attributed to the use of his own HRTFs.

As found in the present study, the use of generic HRTFs generally leads to a significantly higher number of reversals (Wenzel et al., 1993; Bronkhorst, 1995). For example, Wenzel et al. (1993) found that there was a significant increase in front/back and up/down[6] reversals with generic HRTFs. There are some techniques in which a subject's localization performance can be improved when using generic HRTFs. Shinn-Cunningham, Durlach and Held (1998) reported that if subjects were provided with feedback after every trial, then localization on the horizontal plane could be improved. When subjects are allowed to gain experience, i.e., practice, then localizing with generic HRTFs typically does not result in as many reversals or errors (Wenzel et al., 1993). Another method for improving subject performance is to derive the HRTFs from "good" localizers (Wenzel et al., 1993; Middlebrooks, 1999b). Good localizers are subjects whose free-field localization performance is better than average and whose headphone localization performance in virtual auditory space closely matches his/her free-field localization performance. However, F.L. Wightman (personal communication, March 2, 1997) reported that his laboratory has been unsuccessful in documenting any relation between HRTF characteristics and localization performance despite suggestions made in an earlier study (Wightman and Kistler, 1989b). While it is clear that some subjects may have less spectral detail to work with because their pinnae are smooth, it is not clear that this translates into poor performance. With several cues to work with, some individuals seem to emphasize one or more cues depending on their own physical characteristics.

Recently Middlebrooks (1999a, b) investigated an alternative method for improving localization performance when using generic HRTFs. Middlebrooks (1999a) found that by scaling a generic HRTF in frequency to minimize the mismatch between spectral features in the end-listener's and in the individual from whom the HRTF was derived, spectral differences between this pair of subjects could be improved by an average of 15.5%. The optimal scale factor between pairs of subjects correlated highly with the ratio of the subjects' maximum interaural delay, head size and the size of their external pinnae. The penalty for the use of generic HRTFs was reduced by approximately 50% when subjects localize virtual sound sources using a scaled generic HRTF (Middlebrooks, 1999b). If the scaling factor did not closely match the end-listener, then systematic errors would be observed. In particular, if the scaling factor was derived from an individual with a larger head and larger pinnae than the listener, then the subject tended to overshoot the direction of the targeted

---

[5] The measurement of subject RA's own HRTFs was made at the Waisman Centre, University of Wisconsin-Madison, by F.L. Wightman, who's assistance is gratefully acknowledged.

[6] In this paper, the term "up/down" denotes both up-down and down-up reversals.

sound source in the lateral dimension. On the other hand, if the scaling factor was derived from an individual with a smaller head and smaller pinnae than the listener, then the subject tended to undershoot the direction of the targeted sound source in the lateral dimension. In the present study, pinnae and head size for both the individuals from whom the HRTFs were derived and the listeners were not known and thus observations like those of Middlebrooks (1999b) cannot be made. It should be mentioned that while this is a possibility for future consideration, determining the best frequency scale to use in order to maximize performance can be a timely undertaking. In addition, while the scaling technique of Middlebrooks (1999a) appears to improve localization performance when using generic HRTFs, subjects will still make a substantial number of incorrect localization judgements, which is unacceptable in critical mission applications that employ directional auditory cueing.

Free-field localization as a function of two cutoff frequencies

Experiment 3: 3 kHz cutoff frequency

Method

Subjects

Six male and eleven female subjects voluntarily participated in this study. Subjects ranged in age from 15 to 53 years with a mean age of 27.8. Three of the subjects were in-house employees; the others were recruited from the general population outside of DCIEM. One subject (SD) had participated in Experiment 1, one subject (BC) had participated in Experiment 2,and three subjects (KS, RA and SV) had participated in Experiments 1 and 2. In the absence of audiometric equipment, subjects were aurally screened by the experimenter. All subjects reported normal hearing, no recent exposure to loud noises and no history of hearing abnormalities. The use of aural reporting is not uncommon in localization studies (Noble, 1987; Perrott, Sadralodabai, Saberi and Strybel, 1991; Begault and Wenzel, 1993).

Stimuli and apparatus

The stimuli were the same as in Experiment 1. Subjects were individually tested while seated on a chair in a single-wall IAC sound-attenuating listening booth located at the University of Toronto. The ambient level in the booth was less for all frequencies than the maximum allowed for open-ear headphone testing (ANSI-S3.1, 1991). Measured reverberation times in the booth were 0.3 seconds at 125 Hz and 0.2 seconds from 250 Hz to 8 kHz. The subject's chair was positioned in the centre of eight Axiom Millennia loudspeakers (model AX1.2). The frequency response of the loudspeaker was 70-22,000 Hz ±3 dB. The loudspeakers were arranged in a circle with a radius of one meter centred at the seated listener's head. Each loudspeaker was mounted on a Yorkville adjustable microphone stand so that the vertical midpoint of each loudspeaker was at the same height as the listener's ear level. The loudspeakers were placed at 45° intervals ranging from 0° to 315° azimuth, increasing clockwise on the horizontal plane with 0° positioned directly in front of the listener. This loudspeaker array thus coincided with the virtual speaker array used in the VAS conditions. The transmitter of a Polhemus 3Space Fastrak magnetic head tracker was suspended from the ceiling of the sound booth, approximately 20 cm directly above the listener's head to monitor the subject's head position. The tracker's receiver was placed on the top of the listener's head and was held in position by a headband worn by the listener. The booth contained a window and intercom that allowed the experimenter to monitor the subject.

A block diagram of the apparatus is shown in Figure 14. The apparatus consisted of the same setup as in Experiment 1 (Figure 2) with the following exceptions. The TDT PD1 did not spatialize the acoustic stimulus. The Polhemus 3Space Fastrak was routed to the TDT HTI (head tracker interface). The left output from the TDT SW2 drove the left input of the Bryston 2B amplifier. The output from the Bryston 2B amplifier, which served to amplify

the acoustic stimulus, was routed to the TDT PM1-Relay. The TDT PM1-Relay routed the acoustic stimulus to the selected Axiom Millennia loudspeaker via a custom cable harness.

The calibration of the acoustic stimulus was performed with a B&K 4149 ½ inch free-field microphone positioned at the centre of the array of loudspeakers. The height of the microphone corresponded to the height of the centre of the loudspeaker (i.e., the subject's ear level). A 16-second sample of each of the acoustic stimulus conditions presented from the loudspeaker positioned at 0° azimuth made up the signal output. The TDT PA4 was used to set the sound level at 70.2 dB(A) based on the 0° azimuth position. The deviation from the 70.2 dB(A) was ±0.1 dB(A). All other azimuth positions for each of the acoustic stimulus conditions were then attenuated by the same value. The levels emanating from each loudspeaker were within 0.5 dB(A) of each other, based on the subject's head position. Preliminary testing was conducted to ensure that level variations could not be used for loudspeaker identification.

Experimental design and procedure

A 3 (acoustic stimulus) x 8 (azimuth) x 4 (session) within-subject repeated measures design was employed to assess the subjects' ability to localize the acoustic stimulus in the free-field. The acoustic stimulus was spatialized at one of eight static azimuth positions on the horizontal plane (the same azimuth positions used in the VAS conditions in Experiments 1 and 2). A block was comprised of one of the three acoustic stimuli and consisted of 8 practice trials followed by 104 experimental trials. Each azimuth position was presented once in the practice trials and 13 times in the experimental trials. A session contained two repetitions of each block. A Latin Square was employed to counterbalance trials and blocks across subjects and sessions. Following the presentation of one block for each stimulus condition, the subject was given a short break followed by a subsequent presentation of one block for each stimulus. The duration of each session was approximately one hour. Each subject completed four sessions, each on a different day.

The procedure was the same as in Experiments 1 and 2, with the following exception: once a subject was seated on the chair in the booth, he/she was instructed to fixate on a LED located directly in front of him/her on the wall of the booth. If during the presentation of the acoustic stimulus the subject moved his/her head more than two degrees in any direction of yaw, pitch or roll, flashing LEDs notified the subject to reposition his/her head to the "straight-ahead" position. In this instance the trial was discarded and presented again at the end of the current block of trials.

Experiment 4: 4 kHz cutoff frequency

Method

Subjects

Seven male and six female subjects voluntarily participated in this study. Subjects ranged in age from 17 to 40 years with a mean age of 25.7. Three of the subjects were in-house employees. The others were recruited from the general population outside of DCIEM. Five of the subjects (AA, DH, EP, JM, and RT) had participated in Experiment 3, one subject (BC) had participated in Experiments 2 and 3, and two subjects (RA and SV) had

participated in Experiments 1, 2, and 3. In the absence of audiometric equipment, subjects were aurally screened. All subjects reported normal hearing, no recent exposure to loud noises and no history of hearing abnormalities.

Stimuli and apparatus

The 3 kHz cutoff frequency used in Experiments 1-3 was replaced with a 4 kHz cutoff frequency. This was done in order to determine if the frequency of reversals could be reduced. The apparatus of Experiment 3 was used (see Figure 14).

Experimental design and procedure

The experimental design and procedure of Experiment 3 were used.

Results

When subjects localized under free-field conditions, performance was significantly affected by whether the 3 kHz ($F(2,32) = 22.99$, $p < 0.01$) or the 4 kHz ($F(2,24) = 9.73$, $p < 0.01$) cutoff frequency stimulus was low-pass (LP), high-pass (HP) or broadband (BB). With both the 3 kHz and 4 kHz cutoff frequency stimuli, a Scheffe post hoc analysis at the 0.05 alpha level revealed that performance between the BB and HP stimuli was similar. Performance in these stimuli conditions was significantly better than in the LP stimulus condition. As was observed when subjects localized signals under the VAS-quiet and VAS-noise conditions of the present study, performance in free-field conditions was most accurate when subjects were localizing a broadband stimulus. Subject performance for localizing the LP and HP 3 kHz, and LP and HP 4 kHz cutoff frequency stimuli is shown in Figure 15a and 15b, respectively. For the HP stimulus condition, subjects localized more accurately with the 4 kHz than the 3 kHz cutoff frequency stimulus for all azimuth locations, with the exception of 180° (Figure 15b).

Localization performance was significantly better in the HP 4 kHz than the HP 3 kHz as revealed by the data of the eight subjects (AA, BC, DH, EP, JM, RA, RT and SV) who participated in both cutoff conditions ($F(1, 7) = 4.10$, $p < 0.05$). For the LP 3 kHz versus LP 4 kHz cutoff frequency, subject performance was more similar across azimuth positions (Figure 15a). However, as revealed by the data of the eight subjects who participated in both cutoff frequency conditions, localization performance was significantly better in the LP 4 kHz than the LP 3 kHz stimulus condition ($F(1,7) = 28.79$, $p < 0.01$). It is interesting to note that performance was affected more by LP than by HP cutoff frequencies for azimuth locations between 135° and 225°.

There was a significant effect due to sessions when subjects were localizing either a 3 kHz ($F(3,48) = 9.37$, $p < 0.01$) or a 4 kHz ($F(3,36) = 6.87$, $p < 0.01$) cutoff frequency stimulus. A Scheffe post hoc analysis (alpha = 0.05) for both cutoff frequency stimuli showed that subject performance improved over sessions. The effect of azimuth location on performance was only significant when subjects localized the 4 kHz cutoff frequency stimulus ($F(7,84) = 5.43$, $p < 0.01$). However, a Scheffe post hoc analysis at the 0.05 alpha level failed to reveal any significant effect due to azimuth position.

Table 2 shows the distribution of localization judgements classified as corrects, adjusted corrects, errors, and front-back and back-front reversals, as a function of stimulus type under free-field conditions. When adjusted correct performance was examined, only azimuth position was a significant factor in performance when subjects were localizing either the 3 kHz ($F(7,112) = 2.79$, $p < 0.01$) or 4 kHz ($F(7,84) = 6.31$, $p < 0.01$) cutoff frequency stimulus. For localization performance under either the BB or HP stimulus

condition, there was very little difference between the correct and adjusted correct performance for both the 3 kHz and 4 kHz cutoff frequency. However, when subjects localized the LP stimulus there was a larger difference between the correct and adjusted correct performance for the 3 kHz (correct = 77.3%; adjusted correct = 94.5%) as well as the 4 kHz (correct = 85.6%; adjusted correct = 97.9%) cutoff frequency stimuli. There was not a great difference in the number of errors that occurred with either the 3 kHz or 4 kHz cutoff frequency condition under LP, HP or BB conditions. However, there were more back-front reversals for LP conditions with both 3 kHz (13.8%) or 4 kHz (10.5%) stimuli as compared to BB (3 kHz = 0.2%; 4 kHz = 0.2%) and HP (3 kHz = 1.1%: 4 kHz = 1.3%) conditions.

Discussion

Localization performance was poorer in the LP stimulus condition compared to the HP and BB stimuli conditions (Table 3a and b). This appears to be in keeping with the observation that the presence of high frequencies contributes to more accurate and subsequently less variable localization behaviour (Blauert, 1969/70; Wettschurek, 1973; Musicant and Butler, 1984a, b; King and Oldfield, 1997). For example, King and Oldfield (1997) investigated sound localization as a function of the upper and lower limits of signal bandwidth required for accurate localization in a communication system. This was achieved by using progressively lower low-pass cutoff frequencies and higher high-pass cutoff frequencies. These investigators found that elevation errors began to increase in the low-pass condition as the low-pass cutoff frequency approached 9 kHz. Elevation errors began to increase in the high-pass condition as the high-pass cutoff frequency approached 6-9 kHz. This general trend can also be observed in the two cutoff frequency free-field conditions of the present study. The average number of correct localization judgements is 77.3% in the low-pass 3 kHz cutoff frequency condition versus 90.2% in the high-pass 3 kHz cutoff frequency condition. The average number of correct localization judgements is 85.6% in the low-pass 4 kHz cutoff frequency condition versus 94.5% in the high-pass 4 kHz cutoff frequency condition.

General discussion

The experiments described in this paper were run in the order of their presentation. Due to limited access to facilities and subject availability, not all subjects could participate in both VAS and free-field conditions. The localization performance for those subjects who participated in VAS and in free-field are illustrated in Tables 4a, b and c. The assignment of subjects across experiments was not randomly determined due to the aforementioned constraints. Mean localization performance for these subjects was better in the free-field compared to VAS (data are averaged across HRTFs for each stimulus condition).

For those subjects who participated in both the VAS-quiet and free-field 3 kHz cutoff frequency experiments (KS, RA, SD and SV), the average number of reversals for each stimulus condition in VAS is slightly more than doubled compared to the free-field condition (LP: 25.1% versus 11.1%; HP: 28% versus 13.1%; BB: 25.8% versus 12%). However, it should be noted that subject SD had the greatest number of reversals in free-field for all stimulus conditions compared to the other three subjects (see Tables 4a, b and c) and thus her performance affected the mean. Although subject SD had the largest number of reversals in free-field, she did not have the largest number of reversals in VAS. Subject SV had the largest number of reversals in VAS and had the least number of reversals in free-field except in the HP stimulus condition (0.4% versus no reversals by subject KS). For those subjects who participated in both the VAS-noise and free-field 3 kHz cutoff frequency conditions (BC, KS, RA and SV), the average number of reversals for each stimulus condition in VAS is substantially greater than in the free-field (LP: 23.6% versus 6.7%; HP: 20.3% versus 1.5%; BB: 20.8% versus 0.4%). Again, subject SV had the largest number of reversals in the VAS-noise condition and had the least number of reversals in the free-field except in the HP stimulus condition (0.4% versus no reversals by subjects BC and KS). One would expect that the individual with the largest number of reversals in VAS would also have the largest number of reversals in free-field. This observation was indeed the case in the study reported by Wenzel et al. (1993). In that study subject SIH had the fewest reversals in VAS and free-field (10% and 2%, respectively) while subject SIM had the most reversals in VAS and free-field (50% and 43%, respectively).

Higher localization accuracy and fewer reversals in free-field compared to VAS as reported in this paper are not surprising. These results are consistent with the findings of other investigators (Wightman and Kistler, 1989b; Wenzel et al., 1993; Bronkhorst, 1995). Moreover, investigators have also reported greater occurrence of reversals in virtual auditory space regardless of the choice of HRTFs (personal or generic), compared with localization in the free-field. Wightman and Kistler (1989b) found that the percentage of front/back reversals on average when using personal HRTFs was almost twice as high for virtual auditory space as for free-field (11% versus 6%). Wenzel et al. (1993), who performed a similar experiment to Wightman and Kistler (1989b), found that front/back reversals were higher in virtual auditory space than in free-field (31% versus 19%), when generic HRTFs were employed.

In the VAS-quiet experiment, subjects made on average more errors and front/back reversals in the high-pass 3 kHz cutoff frequency condition versus the low-pass 3 kHz cutoff frequency condition. These results appear to be independent of whether generic or personal HRTFs were used. In the two free-field experiments, the average number of errors and front/back reversals were greater in the low-pass cutoff frequency conditions than in the high-pass cutoff frequency conditions. This opposite role was not expected and at present, the authors cannot account for this behavior. Furthermore, the present authors cannot account for the presence of more back-front than front-back reversals in the two free-field experiments in spite of the former occurring less frequently than the latter in other studies (Oldfield and Parker, 1984a, b, 1986). To more fully investigate these differences, it is recommended that a study similar to that of King and Oldfield (1997) be performed for sound sources presented in both VAS and free-field.

One factor that might have contributed to some of the differences in localization performance observed in the present study is the lack of extensive training of and feedback to the subjects. Indeed Wenzel et al. (1993) reported that the amount of experience that the subject has with respect to localization under given experimental conditions can affect performance. Bronkhorst (1995), for example, found that listeners experienced with sound localization made substantially fewer front/back reversals than did inexperienced listeners. Investigators of sound localization differ in their philosophy and technique with respect to the amount of training and feedback given to subjects. Some investigators train subjects extensively with no feedback (e.g., Wightman and Kistler, 1989b). Others train subjects with a stimulus that is not tested in the study with and without visual feedback until subjects' root-mean-square error in degrees reaches a preset criterion (e.g., Middlebrooks, 1992). Yet others (e.g., Wenzel et al., 1993), in addition to the present investigators, give very minimal training and no feedback despite that feedback improves performance (Butler, 1987; Middlebrooks, 1992). Minimal training and no feedback as to the correctness of localization judgements were deliberately given to the subjects in all four experiments in this study because it was important that the subjects' performance reflect that of listeners inexperienced with the localization of a stimulus in VAS or in free-field. It is possible that the user of a 3-D audio display may not have the opportunity for extensive training and/or feedback. However, such limitations may be minimal. For example, Asano et al. (1990), who conducted localization testing in VAS, reported that the number of front/back reversals gradually decreased during the progression of the experiments and reached a steady state in spite of no feedback. These findings are in partial agreement with the present study as subjects experienced a practice effect in Experiments 2-4. This suggests that further testing could have resulted in improved localization accuracy and a decrease in subject variability.

The occurrence of a greater number of front/back reversals in VAS relative to free-field in the present study and that of others (e.g., Wightman and Kistler, 1989b; Wenzel et al., 1993; Bronkhorst, 1995) is not yet fully understood. The method used to measure HRTFs, either with the microphone in the ear canal or at the entrance to the canal, has been identified as a possible source of difference. Another explanation for the discrepancy between the performance of VAS versus free-field sound localization may be that there is an incorrect simulation of high-frequency spectral cues above 7 kHz, arising from a possible distortion introduced by the HRTF measurements performed with probe microphones in the listener's ear canals (Bronkhorst, 1995). As Middlebrooks (1999b) pointed out, the size of the head and pinnae of the individual from whom HRTFs were derived can also affect VAS performance relative to that of free-field. Recently, Martin et al. (submitted) developed and evaluated an HRTF measurement technique. They compared virtual and free-field localization performance across a wide range of sound-source locations for three subjects. For each subject, virtual and free-field localization performance was found to be indistinguishable, as indicated by both front/back reversal rates and average localization errors. The development of a system of such high fidelity is a significant milestone in the maturation of virtual audio technology.

Implications

The present authors acknowledge that the results reported in this preliminary study are partially confounded by an unequal number of subjects in each experiment and the differing measurement resolution used for assessing localization accuracy. Relatively fewer subjects participated in the VAS experiments than in the free-field experiments. This was primarily due to subject availability and limited access to testing facilities. A smaller sample size implies that variability may be affected. For example, in the free-field 3 kHz cutoff frequency condition, there were two subjects (KD and SD) who had substantially lower localization accuracy than the other subjects. Because of the relatively large number of subjects in this experiment, these poorer localizers had less of an effect on variability. On the other hand, with the smaller number of subjects in the VAS experiments, the performance of one subject becomes more critical (e.g., see Figures 10-12 which show extreme subject differences in spite of all subjects using the same HRTF) especially given the measurement resolution that was employed. The measurement resolution of the VAS experiments is coarser

than the measurement resolution of the free-field experiments. In the VAS experiments, each data point is an integer score out of five while in the free-field experiments each data point is an integer score out of 26. If the subject makes one incorrect response on any azimuth position, then his/her score would drop from 100% to 80%. It was desirable to run each subject in the VAS experiments under all seven HRTFs during each session in order to determine if subject proficiency would improve as a function of a given HRTF. Similarly, the present investigators wanted to observe subject performance in the free-field over time given that DCIEM does not have ready access to facilities for testing subjects' localization performance in free-field. In spite of these restrictions, the present data are an initial step towards understanding the impact of some fundamental parameters in sound localization in VAS and free-field. Subsequent studies will address the aforementioned limitations.

Regardless of these limitations, there are several factors for choosing not to implement a 3-D audio system into real-world applications based on the present results. Localization judgements in VAS, expressed as percent correct (Tables 1a and b), were all below 50% in the noise condition regardless of the HRTF or stimulus bandwidth. Localization accuracy in the quiet condition was also below 50%, with the following exceptions: HRTFs "R" and "S" in the LP condition; and HRTFs "R", "S" and "W" in the BB condition. This level of proficiency is clearly unacceptable in a real-world application. Localization performance would still be unacceptably low in spite of a 50% improvement in accuracy using the HRTF scaling technique proposed by Middlebrooks (1999a). It should also be noted that in the present study, subjects were highly trained and visually monitored in the proper fitting and usage of the Stax headphones and the Racal AVH HPD. Neither of these conditions are necessarily characteristic of populations of military personnel who often receive minimal personal instructions and/or vigilant monitoring of the proper care and usage of HPDs. For example, proper fitting of the headphone on the listener's head will ensure that spatialization is not degraded. It is critical that both ear cups are over the pinnae, and that the right ear cup is on the right ear and the left ear cup is on the left ear. It should be noted that the authors have found that it is common practice for operators to remove the headset from one ear to improve direct communication transmission to those nearby. In such instances the presentation of all spatial cues will be disrupted, leading to further degradation in localization accuracy. The present data are also limited to the choice of spatial positions and stimuli. In a general-purpose 3-D audio display, localization accuracy and front/back reversals may depend on the source positions used, the type of stimuli and HRTFs (personal versus generic), and the localization proficiency and experience of the listeners.

Based on the seven generic HRTFs used in the present study, the data suggest that localization accuracy on the horizontal plane is independent of the different techniques used to measure the HRTFs. Regardless of the HRTF measurement technique, the successful implementation of HRTFs in a general-purpose 3-D audio display necessitates that investigators carefully attend to all the intricate steps involved in the measurements of HRTFs. Failure to properly capture the HRTFs could lead to a misrepresentation of the spatial cues, which in turn could cause incorrect localization judgements. For example, F.L. Wightman (personal communication, March 2, 1997) reported deficiencies in the earmold shell described by Wightman and Kistler (1989a), as it changed the resonant frequency of the ear canal. Pralong and Carlile (1994), using a similar measurement approach to Wightman and Kistler (1989a), developed a miniature in-ear recording system to firmly hold the measurement microphone in the subject's ear canal. This minimized internal ear canal reflections. When Pralong and Carlile (1994) compared the localization performance of their HRTFs to those of Wightman and Kistler (1989b), they reported a reduction in the number of front/back reversals in the perception of azimuth and an increase in the accuracy in the judgement of elevation. These differences were not observed in the present study primarily due to testing only on the horizontal plane.

Conclusions and recommendations

The present research was undertaken to investigate localization performance in VAS on the horizontal plane, in quiet and in the presence of diffuse ambient Leopard tank noise as a function of generic HRTFs and signal bandwidth. The Leopard tank noise was chosen to represent a real-world military sound environment. Localization reversals were not corrected. The outcome of this preliminary study, based on the limited data collected, revealed that localization accuracy, as measured by average percent correct and front/back reversals, was higher in the two free-field experiments than in the two VAS experiments. Subject performance in VAS was not significantly affected by type of generic HRTF; localization accuracy using the broadband stimulus was not significantly better than the low-pass 3 kHz stimulus. This finding suggests that the role of spectral cues is minimal for sound sources located on the horizontal plane and implies that the restriction of the bandwidth of the communication system to 3.5 kHz might not significantly impede user localization accuracy in VAS. In the presence of diffuse ambient Leopard tank noise localization performance was degraded compared to the quiet condition, suggesting that 3-D audio technology may not yet be very useful in present-day military noisy listening environments. Average localization performance in the free-field low- and high-pass 4 kHz cutoff frequency conditions was slightly better than the free-field low- and high-pass 3 kHz cutoff frequency conditions. Given this latter result, it is assumed that this improved level in performance would also be observed in VAS.

At present, several critical factors are impeding the transition potential of 3-D audio technology into real-world applications. These include the choice of personal versus generic HRTFs, environmental factors and communication bandwidth. In particular, performance in VAS is more accurate and results in fewer localization reversals with personal HRTFs compared to generic ones. However, personal HRTFs are traditionally derived from binaural measurements in the ears of the end-listener seated in an anechoic chamber. This requires a substantial investment in infrastructure and equipment, and is presently impractical in most applications. The work of Middlebrooks (1999a) needs to be further developed to quickly and accurately select and/or modify a generic HRTF for the targeted application. The effect of diffuse ambient noise on user performance with either personal or generic HRTFs also requires further investigation. The hardware limitation imposed on the communication bandwidth needs to be addressed particularly when virtual sound sources are presented off the horizontal plane. Until the above issues are more fully understood and resolved it may be prudent to proceed cautiously before the adoption of a 3-D audio system into critical mission applications. Advances in these issues are being made as demonstrated by Martin et al. (submitted). Their results suggest that the imperfection in the simulation of virtual sound sources is an obstacle that may have been surmounted. This represents a significant achievement in the reproduction of spatial synthesis.

Further research is required in the following areas:

- Personal versus generic HRTFs and equalization of the headphones are issues to be considered;
- Testing off the horizontal plane needs to be investigated;
- Techniques for minimizing reversals need to be developed;
- The effects of mismatched pinnae and head size between the end-listener and the individual from whom the HRTF was derived on localization performance needs to be reduced;

- Effects of the hardware bandwidth limitation imposed by existing communication systems need to be understood particularly when virtual sound sources are presented off the horizontal plane;
- The frequency response of the headphone transducer needs to be taken into consideration;
- The level and frequency spectrum of the diffuse ambient noise on auditory localization performance in VAS needs to be investigated;
- Hearing protection devices with a good seal need to be developed to prevent distortion of spatial cues in VAS;

References

1. American National Standards Institute. (1991). **Maximum Permissible Ambient Noise Levels for Audiometric Test Rooms**. (ANSI-S3.1). New York.

2. Arrabito, G.R., Cheung, B., Crabtree, R.B. and McFadden, S.M. (2000). **Release from masking in virtual auditory space during sustained positive acceleration**. *Aviat. Space Environ. Med.*, 71(5), 536-542.

3. Asano, F., Suzuki, Y. and Sone, T. (1990). **Role of spectral cues in median plane localization**. *J. Acoust. Soc. Am.*, 88(1), 159-168.

4. Begault, D.R. (1993). **Head-up auditory displays for traffic collision avoidance system advisories: A preliminary investigation**. *Hum. Factors*, 35(4), 707-717.

5. Begault, D. and Wenzel, E. (1993). **Headphone localization of speech**. *Hum. Factors*, 35(2), 361-376.

6. Begault, D.R. and Erbe, T. (1994). **Multichannel spatial auditory display for speech communications**. *J. Aud. Eng. Soc.*, 42(10), 819-826.

7. Begault, D.R. and Pittman, M.T. (1996). **Three-dimensional audio versus head-down traffic alert and collision avoidance system displays**. *Int. J. Aviat. Psych.*, 6(1), 79-93.

8. Belendiuk, K. and Butler, R.A. (1975). **Monaural localization of low-pass noise bands in the horizontal plane**. *J. Acoust. Soc. Am.*, 58(3), 701-705.

9. Blauert, J. (1983). **Spatial Hearing: The Psychophysics of Human Sound Localization**. J. Allen, Trans. Cambridge, MA: MIT Press.

10. Blauert, J. (1969/70). **Sound localization in the median plane**. *Acustica*, 22(4), 205-213.

11. Bloom, P.J. (1977a). **Creating source elevation illusions by spectral manipulation**. *J. Aud. Eng. Soc.*, 25(9), 560-565.

12. Bloom, P.J. (1977b). **Determination of monaural sensitivity changes due to the pinna by use of minimum-audible-field measurements in the lateral vertical plane**. *J. Acoust. Soc. Am.*, 61(3), 820-828.

13. Bronkhorst, A.W. (1995). **Localization of real and virtual sound sources**. *J. Acoust. Soc. Am.*, 98(5), 2542-2553.

14. Bronkhorst, A.W., Veltman, J.A.H. and van Breda, L. (1996). **Application of a three-dimensional auditory display in a flight task**. *Hum. Factors*, 38(1), 23-33.

15. Burkhard, M. and Sachs, R. (1975). **Anthropometric manikin for acoustic research**. *J. Acoust. Soc. Am.*, 58(1), 214-222.

16. Butler, R.A. (1987). **An analysis of monaural displacement of sounds in space**. *Percept. Psychophys.*, 41(1), 1-7.

17. Butler, R.A. (1986). **The bandwidth effect on monaural and binaural localization**. *Hearing Res.*, 21, 67-73.

18. Butler, R.A. and Belendiuk, K. (1977). **Spectral cues utilized in the localization of sound in the median sagittal plane**. *J. Acoust. Soc. Am.*, 61(5), 1264-1269.

19. Butler, R.A. and Flannery, R. (1980). **The spatial attributes of stimulus frequency and their role in monaural localization of sound in the horizontal plane**. *Percept. Psychophys.*, 28(5), 449-457.

20. Butler, R.A. and Planert, N. (1976). **The influence of stimulus bandwidth on localization of sound in space**. *Percept. Psychophys.*, 19(1), 103-108.

21. Calhoun, G.L., Janson, W.P. and Valencia, G. (1988). **Effectiveness of three-dimensional auditory directional cues**. In: *Proceedings of the Human Factors Society 32nd Annual Meeting*, (pp. 68-72). Santa Montica, CA: Human Factor Society.

22. Calhoun, G.L., Valencia, G. and Furness, T.A. (1987). **Three-dimensional auditory cue simulation for crew station design/evaluation**. In *Proceedings of the Human Factors Society 31st Annual Meeting*, (pp. 1398-1402). Santa Montica, CA: Human Factor Society.

23. Carlile, S. and Pralong, D. (1994). **The location-dependent nature of perceptually salient features of the human head-related transfer functions**. *J. Acoust. Soc. Am.*, 95(6), 3445-3459.

24. D'Angelo, W.R., Bolia, R.S., McKinley, R.L. and Perrott, D.R. (1997). **Auditory guided visual search with visual distractors**. In: *Proceedings of the Acoustical Society of America's 133rd Meeting*, 101(5), (p. 3107). State College, PA: Acoustical Society of America.

25. Doll, T.J., Gerth, J.M., Engelman, W.R. and Folds, D.J. (1986). **Development of Simulated Directional Audio for Cockpit Applications**. (Report no. AAMRL-TR-86-014). Ohio: Wright-Patterson Airforce Base, Armstrong Aerospace Medical Research Laboratory.

26. Durlach, N.I., Rigopulos, A., Pang, X.D., Woods, W.S., Kulkarni, A., Colburn, H.S. and Wenzel, E.M. (1992). **On the externalization of auditory images**. *Presence*, 1(2), 251-257.

27. Ericson, M.A. and McKinley, R.L. (1997). **The intelligibility of multiple talkers separated spatially in noise**. In: R.H. Gilkey and T.R. Anderson (Eds.), *Binaural and Spatial Hearing in Real and Virtual Environments*, (pp. 701-724). New Jersey: Lawrence Erlbaum Associates.

28. Flannery, R. and Butler, R.A. (1981). **Spectral cues provided by the pinna for monaural localization in the horizontal plane**. *Percept. Psychophys.*, 29(5), 438-444.

29. Forshaw, S.E. and Crabtree, R.B. (1983). **Noise and Vibration Assessment of the Leopard C1 Main Battle Tank**. DCIEM No. 83-R-39.

30. Furness, T.A. (1986). **The super cockpit and its human factors challenges**. In: *Proceedings of the Human Factors Society 30th Annual Meeting*, (pp. 48-52). Dayton, OH.

31. Gardner, B. and Martin, K. (1994). **HRTF Measurements of a KEMAR Dummy-Head Microphone**. MIT Media Lab.

32. Good, M.D., Gilkey, R.H. and Ball, J.M. (1997). **The relation between detection in noise and localization in noise in the free-field**. In R.H. Gilkey and T.R. Anderson (Eds.), *Binaural and Spatial Hearing in Real and Virtual Environments*, (pp. 683-699). New Jersey: Lawrence Erlbaum Associates.

33. Green, D.M. and Swets, J.A. (1966). **Signal Detection Theory and Psychophysics**. John Wiley and Sons, Inc: New York.

34. Hebrank, J. and Wright, D. (1974). **Spectral cues used in the localization of sound sources on the median plane**. *J. Acoust. Soc. Am.*, 56(6), 1829-1834.

35. Ivarsson, C., de Ribaupierre, Y. and de Ribaupierre, F. (1980). **Functional ear asymmetry in vertical localization**. *Hearing Res.*, 3, 241-247.

36. King, R.B. and Oldfield, S.R. (1997). **The impact of signal bandwidth on auditory localization: Implications for the design of 3-dimensional audio displays**. *Hum. Factors*, 39(2), 287-295.

37. Leong, P., Tucker, T. and Carlile, S. (1996). **Digital signal processing for the auditory scientist: A tutorial introduction**. In: S. Carlile (Ed.), *Virtual auditory space: Generation and applications*. Landes: Austin, 79-108.

38. Martin, R.L., McAnally, K.I. and Senova, M. (submitted). **Free-field equivalent localisation of virtual audio**. *J. Audio Eng. Soc.*

39. McKinley, R.L. and Ericson, M.A. (1997). **Flight demonstration of a 3-D auditory display**. In: R.H. Gilkey and T.R. Anderson (Eds.), *Binaural and Spatial Hearing in Real and Virtual Environments*, (pp. 683-699). New Jersey: Lawrence Erlbaum Associates.

40. Middlebrooks, J.C. (1999a). **Individual differences in external-ear transfer functions reduced by scaling in frequency**. *J. Acoust. Soc. Am.*, 106(3), 1480-1492.

41. Middlebrooks, J.C. (1999b). **Virtual localization improved by scaling nonindividualized external-ear transfer functions in frequency**. *J. Acoust. Soc. Am.*, 106(3), 1493-1510.

42. Middlebrooks, J.C. (1992). **Narrow-band sound localization related to external ear sources.** *J. Acoust. Soc. Am.*, 92(5), 2607-2624.

43. Mills, A.W. (1972). **Auditory localization**. In: J.V. Tobias (Ed.), *Foundations of Modern Auditory Theory Vol. II*, (pp. 301-348). New York: Academic Press.

44. Moller, H. (1992). **Fundamentals of binaural technology**. *Appl. Acoust.*, 36, 171-218.

45. Moller, H., Sorenson, M.F., Jensen, C.B. and Hammershoi, D. (1996). **Binaural technique: Do we need individual recordings?** *J. Audio Eng. Soc.*, 44(6), 451-469.

46. Moller, H., Sorensen, M.F., Hammershoi, D. and Jensen, C.B. (1995). **Head-related transfer functions of human subjects**. *J. Aud. Eng. Soc.*, 43(5), 300-321.

47. Musicant, A.D. and Butler, R.A. (1985). **Influence of monaural spectral cues on binaural localization**. *J. Acoust. Soc. Am.*, 77(1), 202-208.

48. Musicant, A.D. and Butler, R.A. (1984a). **The influence of pinnae-based spectral cues on sound localization**. *J. Acoust. Soc. Am.*, 75(4), 1195-1200.

49. Musicant, A.D. and Butler, R.A. (1984b). **The psychophysical basis of monaural localization.** *Hearing Res.*, 14(2), 185-190.

50. Nixon, M.S., Anderson, T., Morris, L., McCavitt, A, McKinley, R., Yeager, D. and McDaniel, M. (1998). **Female voice communications in high level aircraft cockpit noises - Part II: Vocoder and automatic speech recognition systems.** *Aviat. Space Environ. Med.*, 69(11), 1087-1094.

51. Noble, W. (1987). **Auditory localization in the vertical plane: Accuracy and constraint on bodily movement.** *J. Acoust. Soc. Am.*, 82(5), 1631-1636.

52. Oldfield, S.R. and Parker, S.P.A. (1986). **Acuity of sound localisation: A topography of auditory space. 3. Monaural hearing conditions.** *Percept.*, 15(1), 67-81.

53. Oldfield, S.R. and Parker, S.P.A. (1984a). **Acuity of sound localization: A topography of auditory space. 1. Normal hearing conditions.** *Percept.*, 13(5), 581-600.

54. Oldfield, S.R. and Parker, S.P.A. (1984b). **Acuity of sound localization: A topography of auditory space. 2. Pinna cue absent.** *Percept.*, 13(5), 601-617.

55. Oppenheim, A.V. and Schafer, R.W. (1989). **Discrete-Time Signal Processing.** Prentice-Hall, Inc., Englewood Cliffs, New Jersey, U.S.A.

56. Patterson, R. (1982). **Guidelines for Auditory Warning Systems on Civil Aircraft.** London, UK: Civil Aviation Authority. Report No.: CAA Paper 82017.

57. Perrott, D.R., Cisneros, J., McKinley, R.L. and D'Angelo, W.R. (1996). **Aurally aided visual search under virtual and free-field listening conditions.** *Hum. Factors*, 38(4), 702-715.

58. Perrott, D.R., Sadralodabai, T., Saberi, K. and Strybel, T.Z. (1991). **Aurally aided visual search in the central visual field: Effects of visual load and visual enhancement of the target.** *Hum. Factors*, 33(4), 389-400.

59. Plenge, G. (1974). **On the difference between localization and lateralization.** *J. Acoust. Soc. Am.*, 56(3), 944-951.

60. Pralong, D. and Carlile, S. (1994). **Measuring the human head-related transfer functions: A novel method for the construction and calibration of a miniature "in-ear" recording system.** *J. Acoust. Soc. Am.*, 95(6), 3435-3444.

61. Ricard, G.L. and Meirs, S.L. (1994). **Intelligibility and localization of speech from virtual directions.** *Hum. Factors*, 36(1), 120-128.

62. Shaw, E.A.G. (1975). **The external ear: New knowledge. Ear moulds and associated problems.** In: *7th Danavox Symposium*, Gl. Avernes, Denmark.

63. Shaw, E.A.G. (1974). **The external ear.** In: W.D. Keidel and W.D. Neff (Eds.), *Handbook of Sensory Physiology*, 5, (pp. 455-490). Berlin-New York: Springer-Verlag.

64. Shaw, E.A.G. (1965). **Earcanal pressure generated by a free sound field.** *J. Acoust. Soc. Am.*, 39(3), 465-470.

65. Shinn-Cunningham, B.G., Durlach, N.I. and Held, R.I. (1998). **Adapting to supernormal auditory localization cues: I. Bias and resolution.** *J. Acoust. Soc. Am.*, 103(6), 3656-3666.

66. Stinnett, T.A. (1989). **Human factors in the supercockpit.** In: R. Jenson (ed.), *Aviat. Psych.*, (pp. 1-37). Brookfield, VE: Gower Publishing.

67. Watkins, A.J. (1978). **Psychoacoustical aspects of synthesized vertical locale cues.** *J. Acoust. Soc. Am.*, 63(4), 1152-1165.

68. Wenzel, E.M. (1994). **Spatial sound and sonification.** In: G. Kramer (ed.), *Auditory Display Sonification, Audification, and Auditory Interfaces*, (pp. 127-150). Reading, MA: Westley Publishing Co.

69. Wenzel, E.M., Arruda, M., Kistler, D.J. and Wightman, F.L. (1993). **Localization using non-individualized head-related transfer functions.** *J. Acoust. Soc. Am.*, 94(1), 111-123.

70. Wettschurek, R. (1973). **The absolute difference threshold of directional perception in the median plane in natural hearing and in hearing over a dummy-head reproducing system.** *Acustica*, 28(4), 197-208.

71. Wightman, F.L. and Kistler, D.J. (1989a). **Headphone simulation of free-field listening. Part I: Stimulus synthesis.** *J. Acoust. Soc. Am.*, 85(2), 858-867.

72. Wightman, F.L. and Kistler, D.J. (1989b). **Headphone simulation of free-field listening. Part II: Psychophysical validation.** *J. Acoust. Soc. Am.*, 85(2), 868-878.

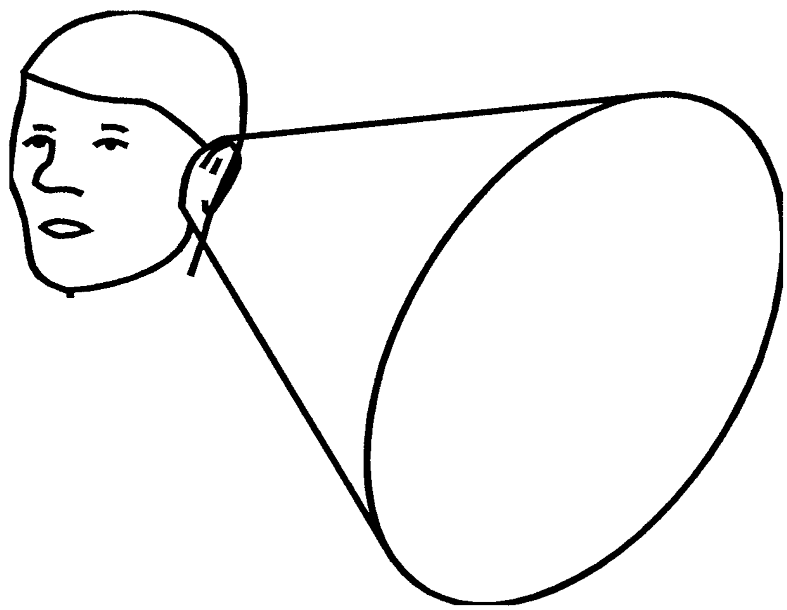

Figure 1: A cone of confusion for a spherical head and a particular interaural time delay. All sound sources on the surface of the cone would produce that interaural time delay. For details of how to calculate the cone of confusion see Mills (1972). Reprinted with permission from Moore BCJ. An Introduction to the Psychology of Hearing. London: Academic Press, 1989.

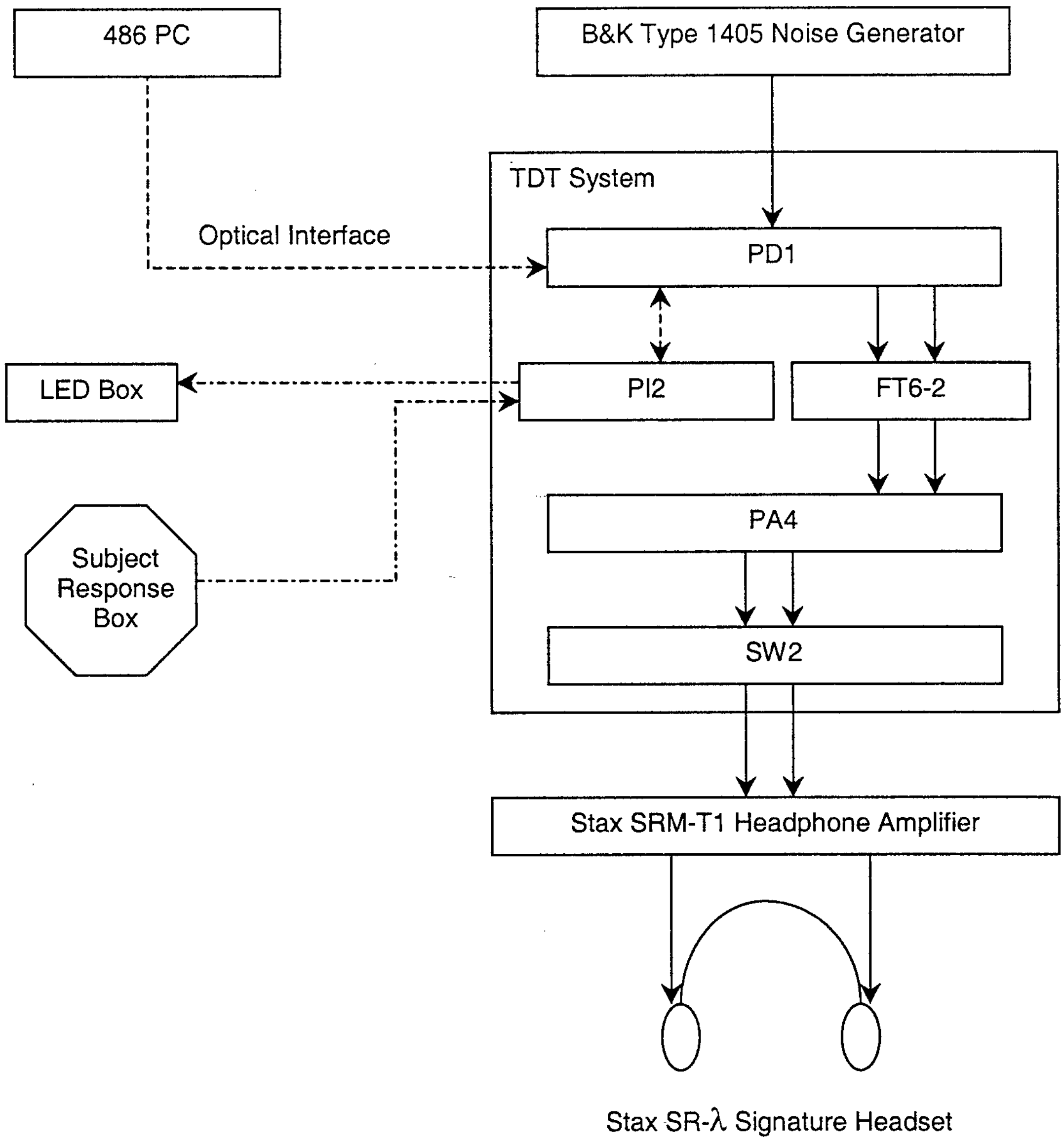

Figure 2: Hardware configuration for Experiments 1 and 2.

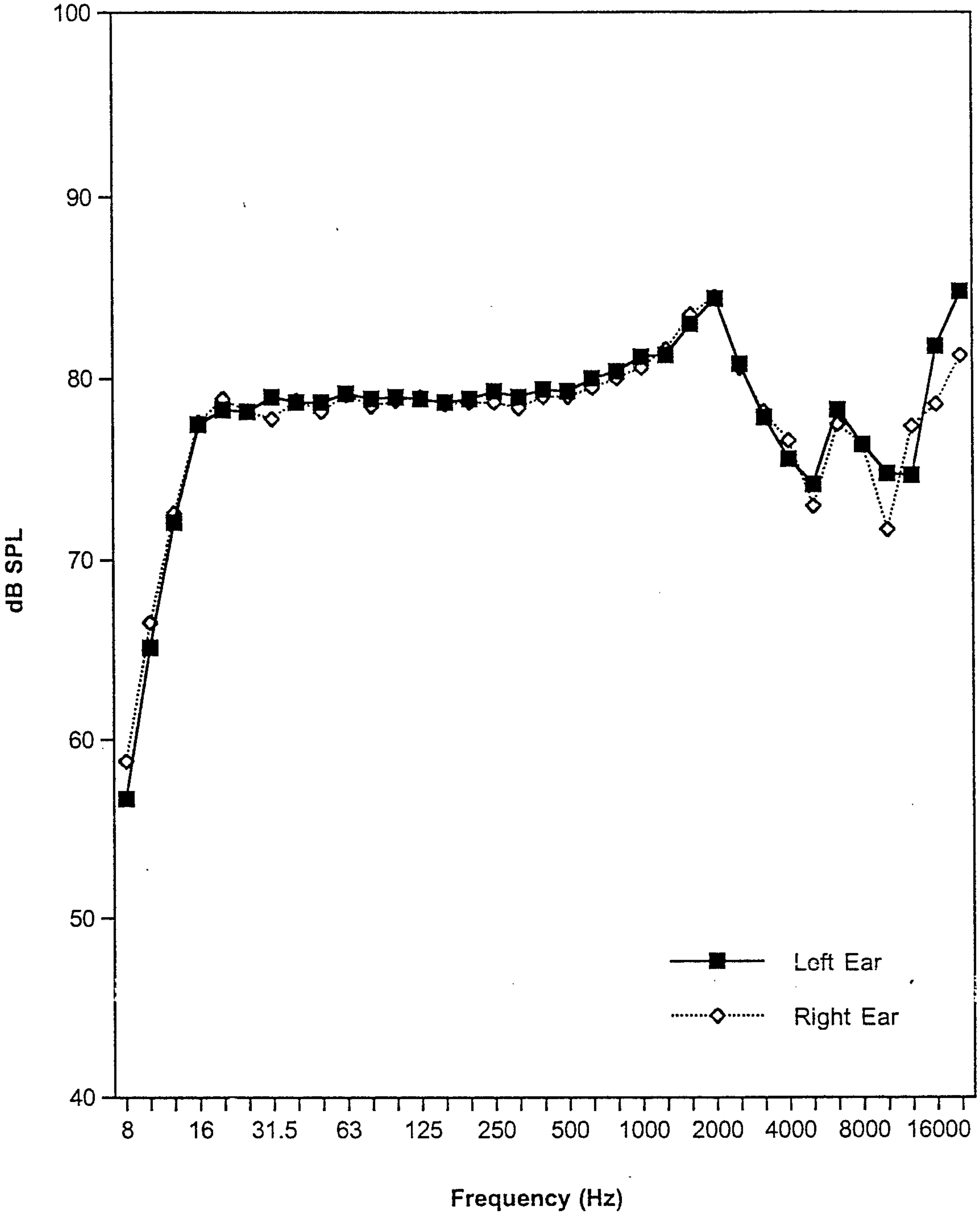

Figure 3: STAX headset frequency response. Differences between left and right earcup output levels against a flat-plate coupler for a STAX (SR- λ) headset, showing negligible divergence. Measurements were made using a pink noise signal source set to produce a nominal level of 85 dB(A) at the coupler.

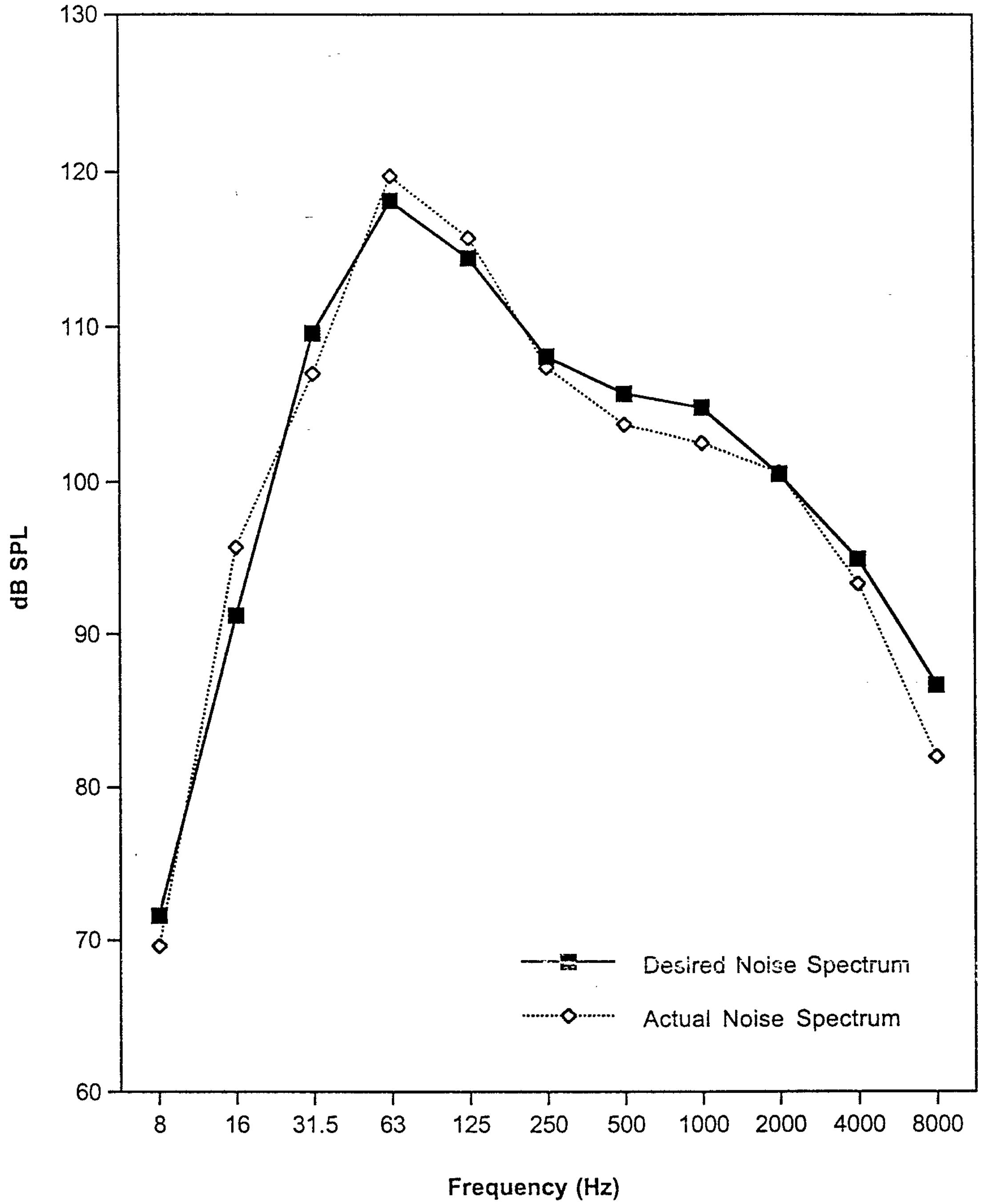

Figure 4: Desired Leopard tank noise spectrum versus achieved noise spectrum. The differences in spectra in actual Leopard noise, reproduced in DCIEM's Noise Simulation Facility.

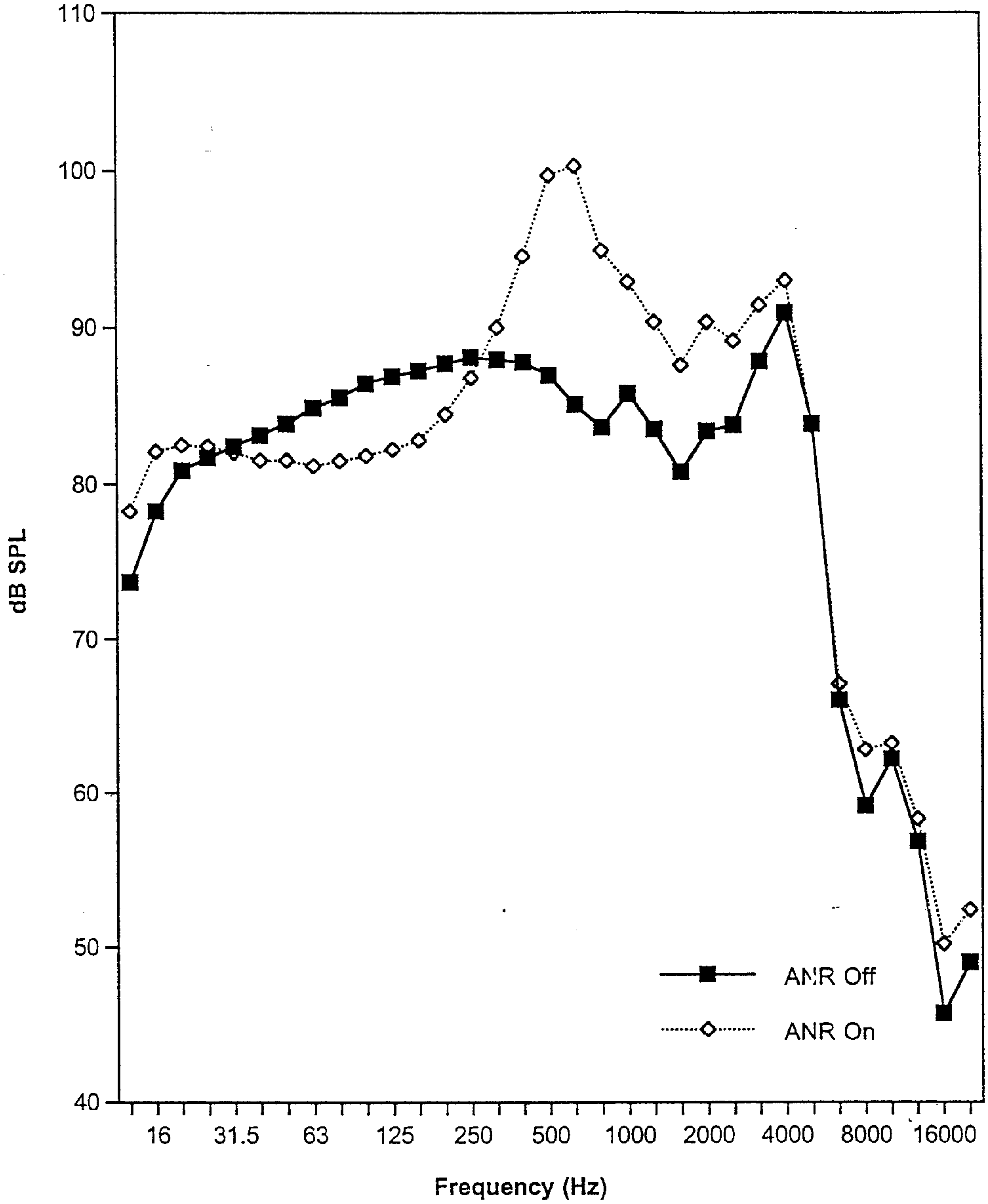

Figure 5a: Frequency response of the Racal Armored Vehicle headset, measured against a flat-plate coupler with Active Noise Reduction (ANR) in active and passive modes. Measurements were made using a pink noise signal source set to produce a nominal level of 85 dB(A) at the coupler with ANR in passive mode.

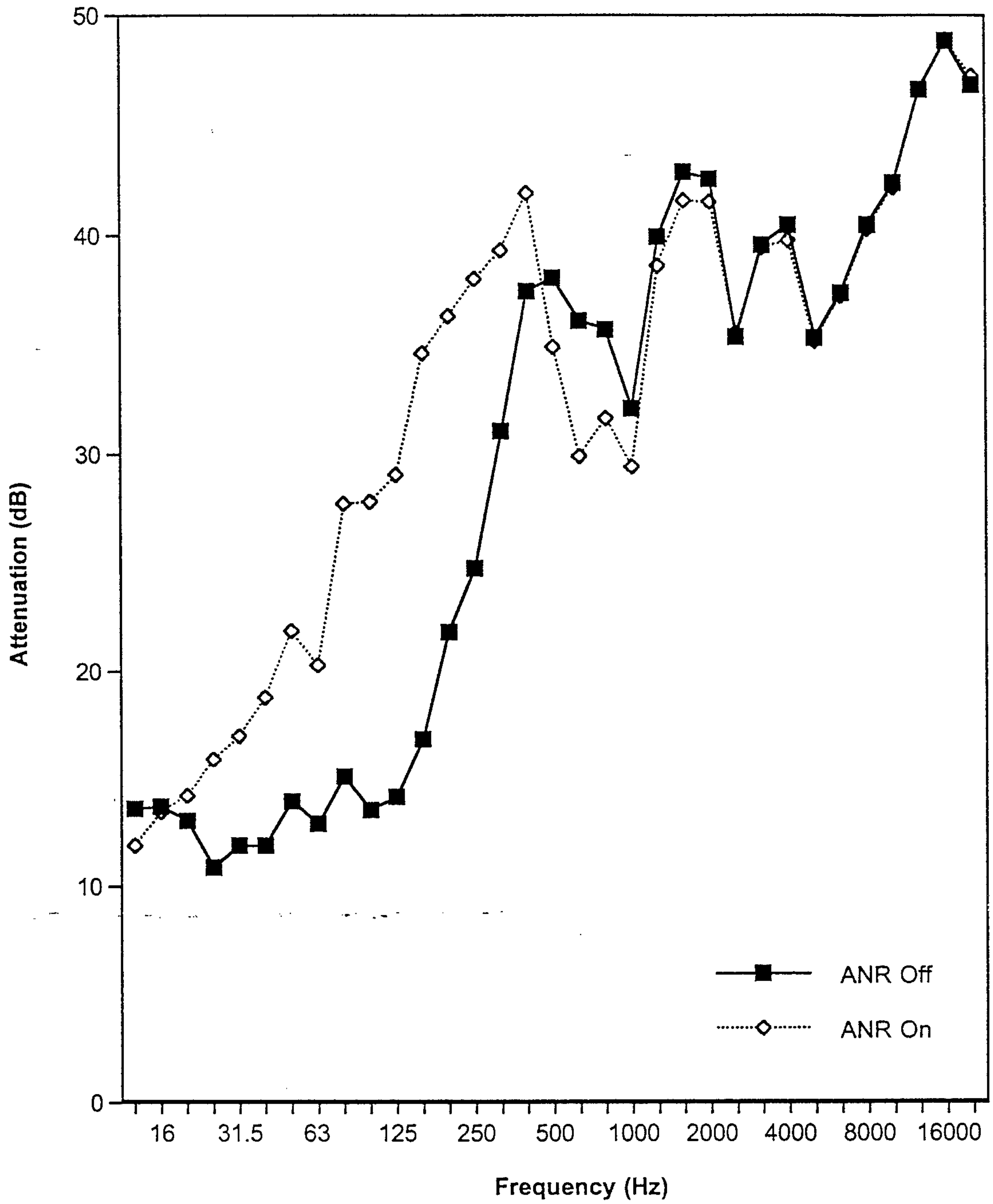

Figure 5b: Sound Attenuation (insertion loss) achieved by the Racal Armored Vehicle headset with an airtight seal against a flat-plate coupler in both active and passive modes. Upon fitting to humans, one would expect some decrement in sound attenuation performance due to fitting anomalies.

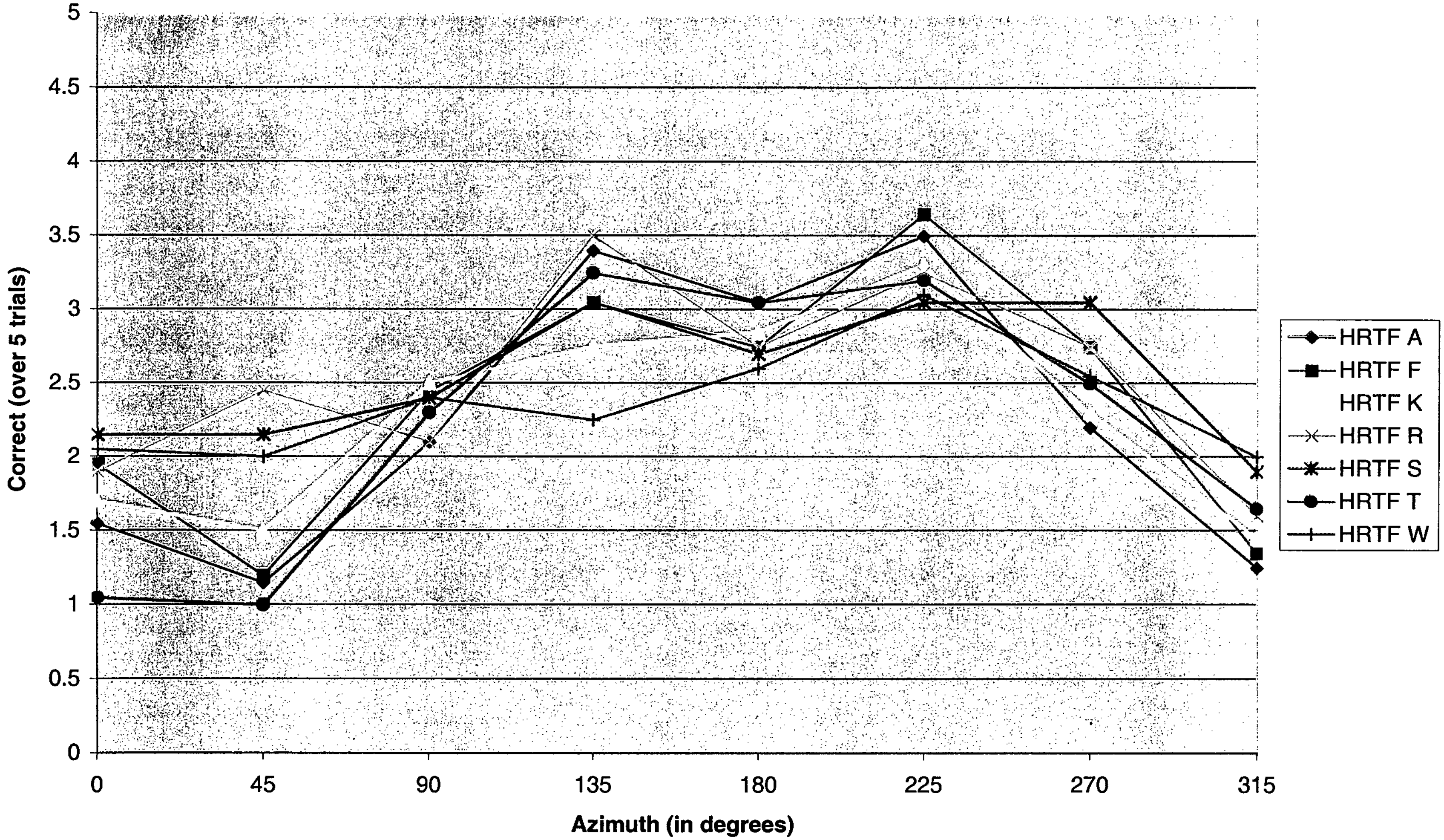

Figure 6a: Correct localization judgements in the low-pass 3 kHz condition of Experiment 1 (VAS - quiet) as a function of HRTFs. Data are averaged across 5 subjects and 4 sessions.

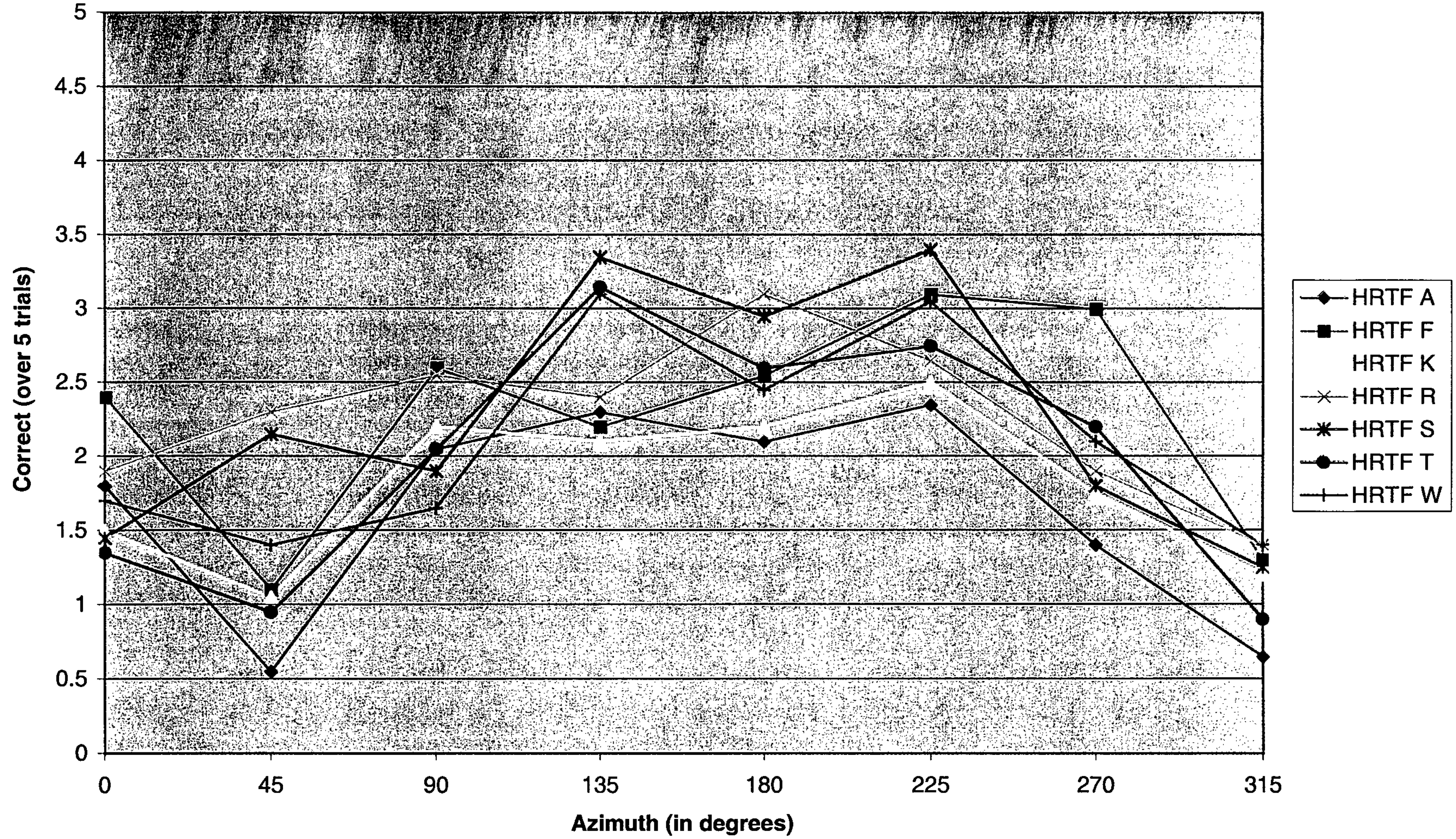

Figure 6b: Correct localization judgements in the high-pass 3 kHz condition of Experiment 1 (VAS - quiet) as a function of HRTFs. Data are averaged across 5 subjects and 4 sessions.

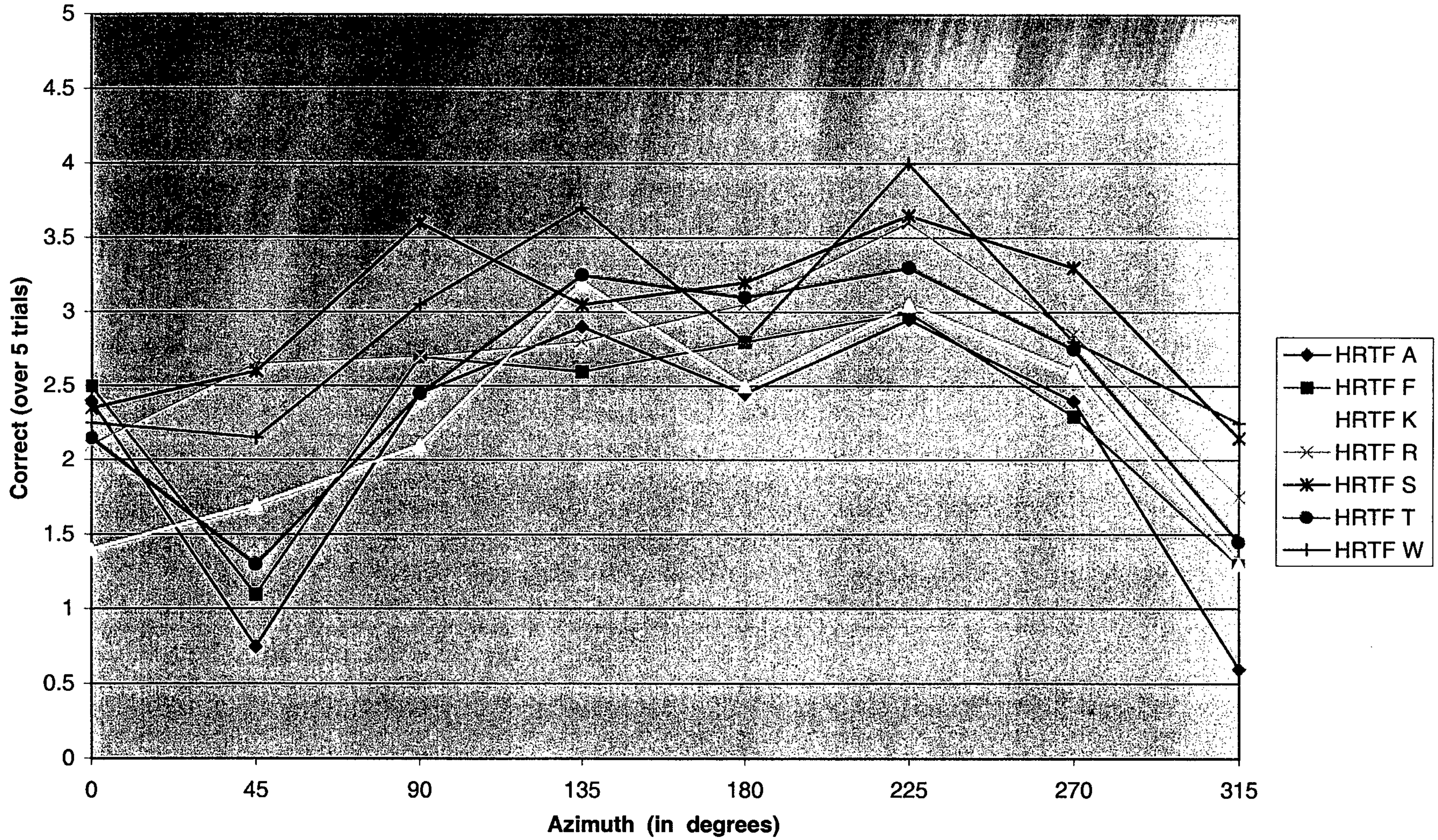

Figure 6c: Correct localization judgements in the broadband condition of Experiment 1 (VAS - quiet) as a function of HRTFs. Data are averaged across 5 subjects and 4 sessions.

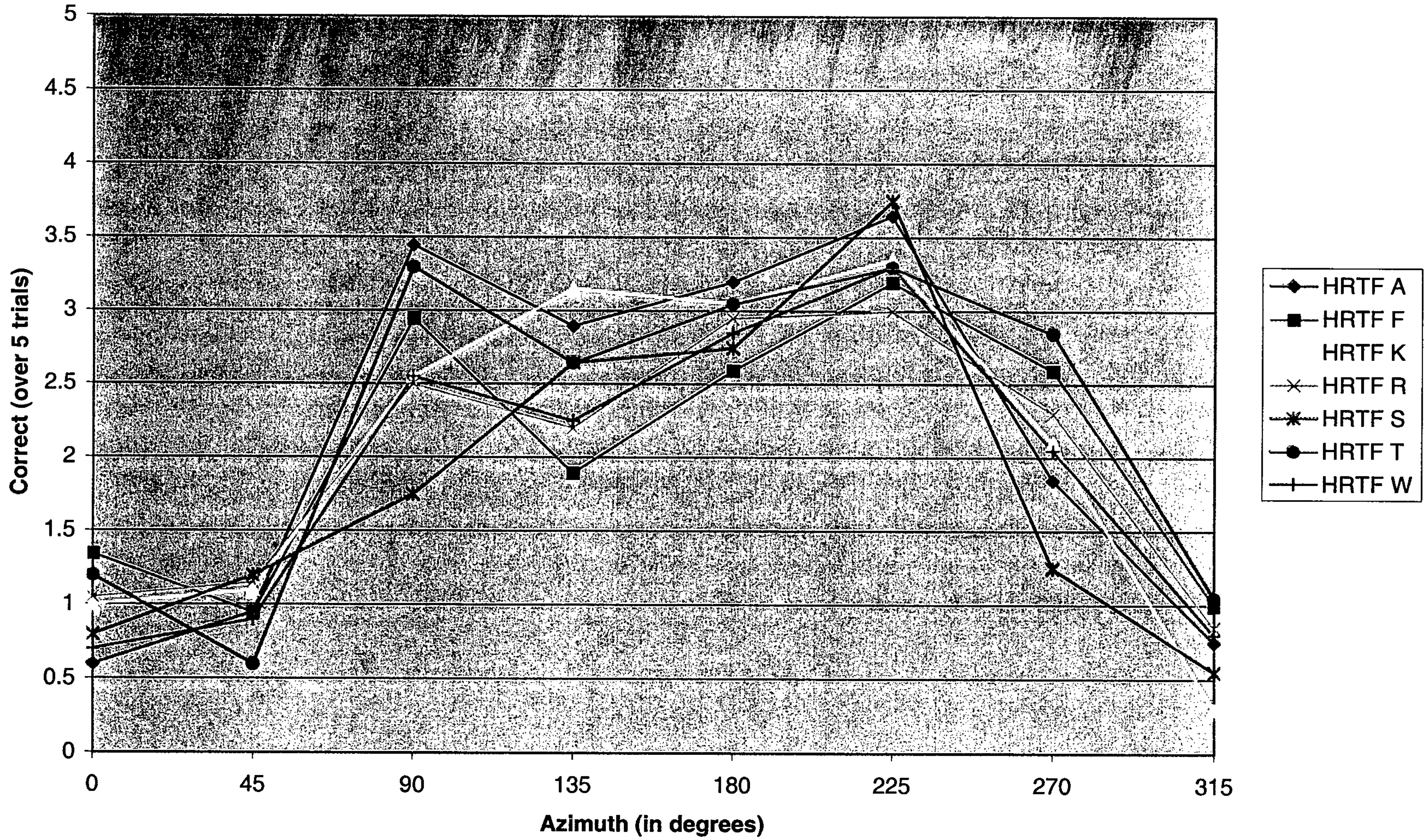

Figure 7a: Correct localization judgements in the low-pass 3 kHz condition of Experiment 2 (VAS - noise) as a function of HRTFs. Data are averaged across 5 subjects and 4 sessions.

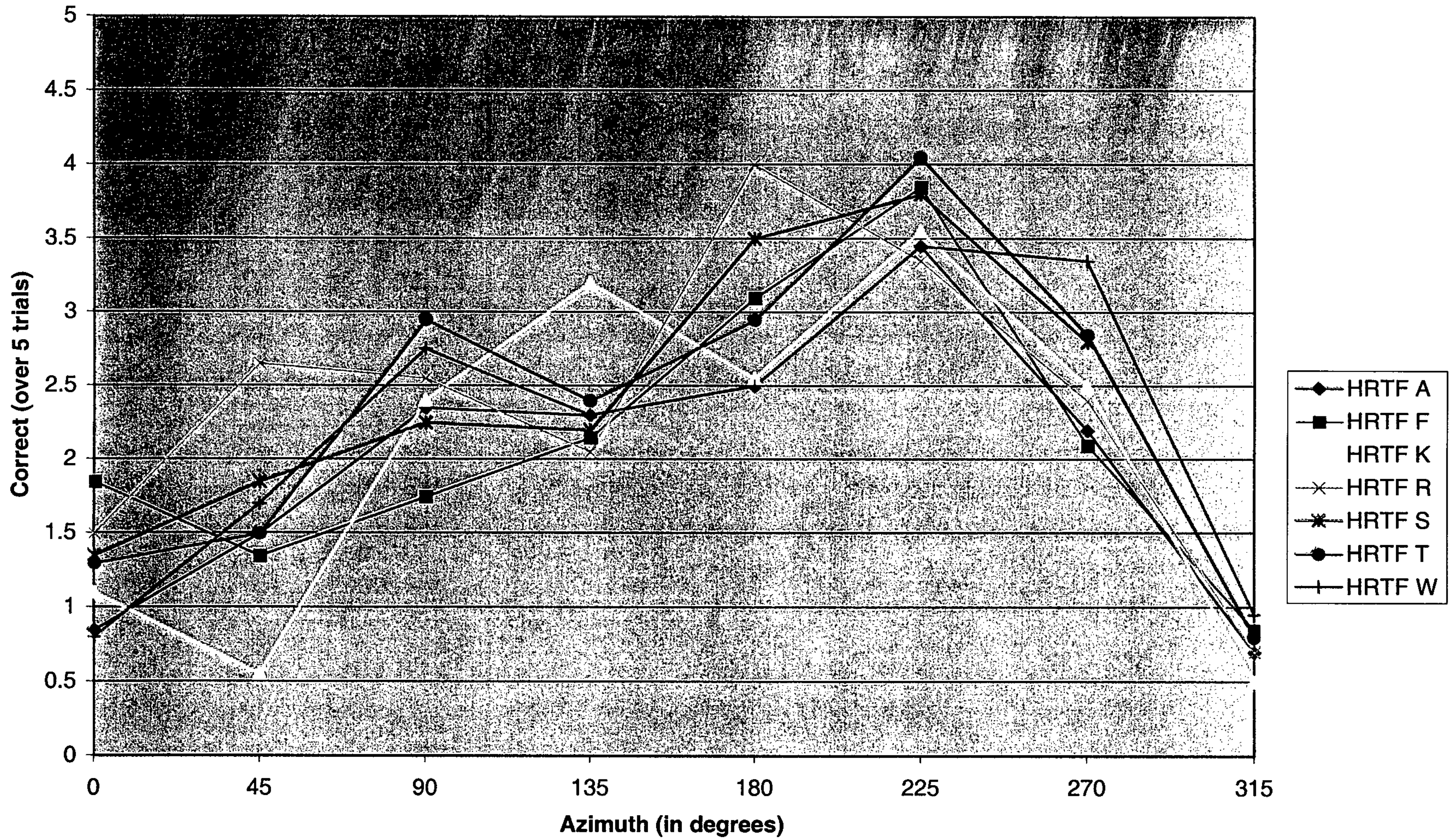

Figure 7b: Correct localization judgements in the high-pass 3 kHz condition of Experiment 2 (VAS - noise) as a function of HRTFs. Data are averaged across 5 subjects and 4 sessions.

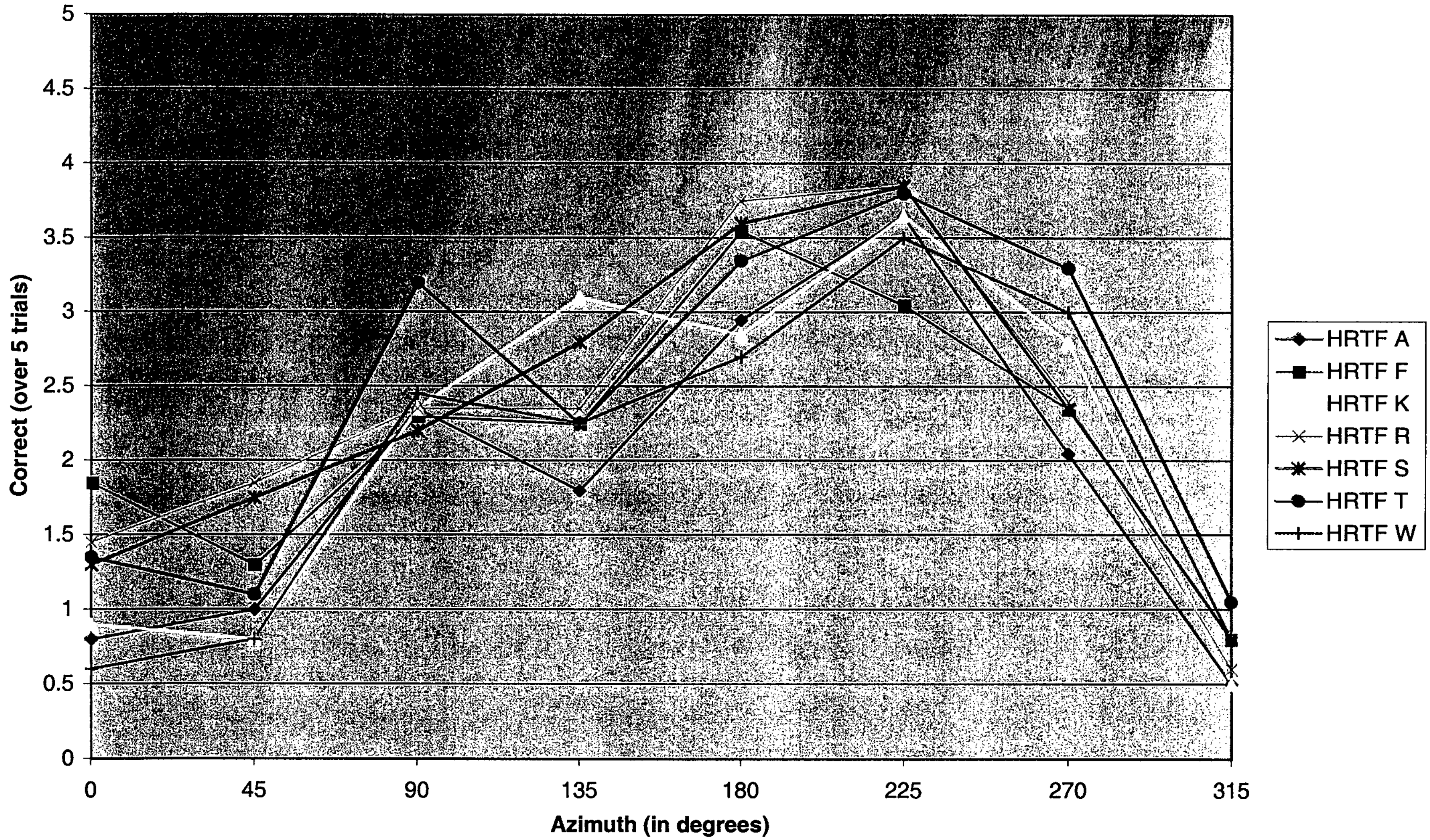

Figure 7c: Correct localization judgements in the broadband condition of Experiment 2 (VAS - noise) as a function of HRTFs. Data are averaged across 5 subjects and 4 sessions.

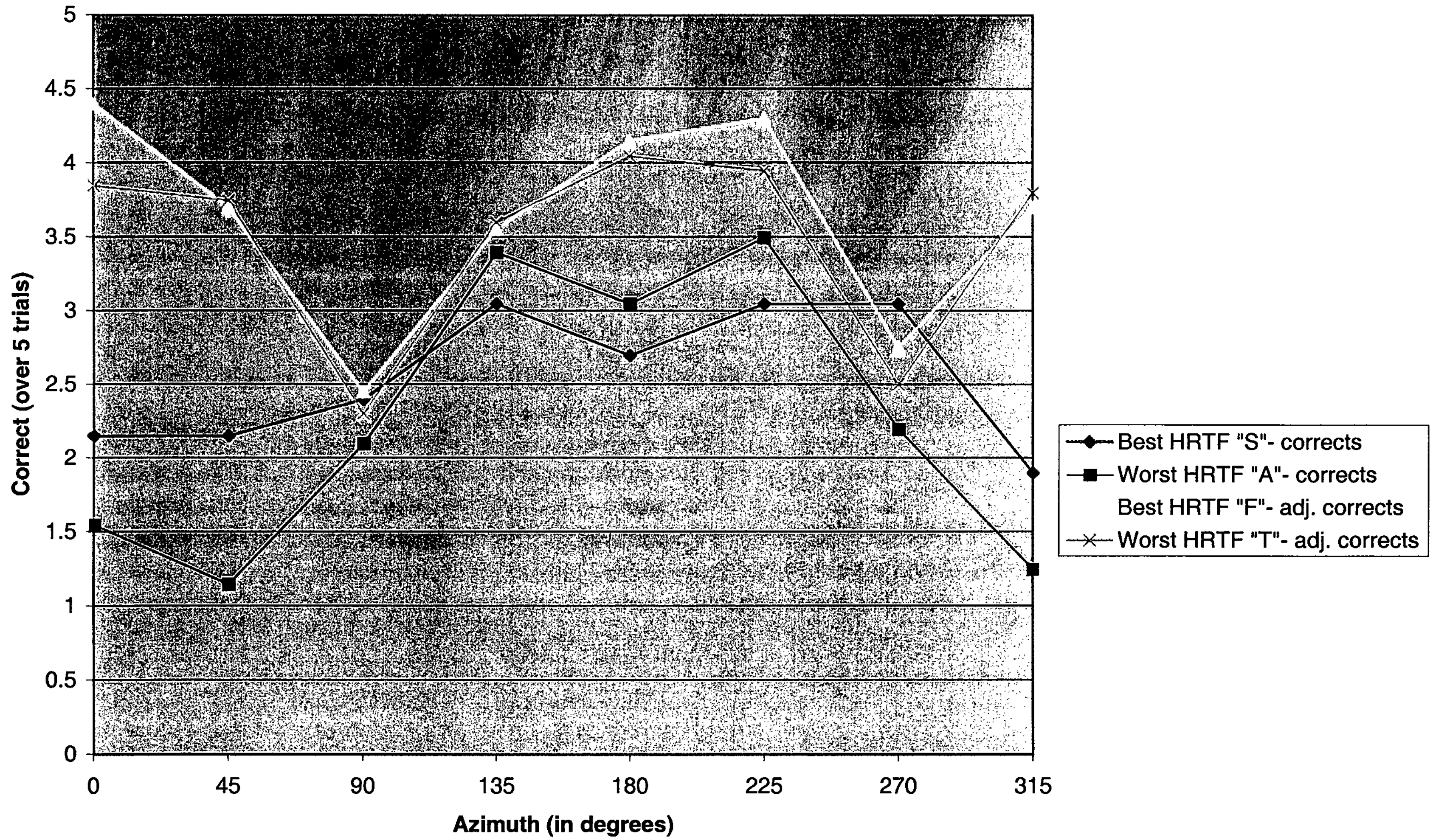

Figure 8a: A comparison of HRTFs that yielded best and worst localization accuracy for corrects and adjusted corrects in the low-pass 3 kHz condition of Experiment 1 (VAS - quiet). Differentiation is made between "corrects" and "adjusted corrects." Adjusted corrects resolve reversals by coding the subjects' responses as if they were made in the correct hemisphere. Data are averaged across 5 subjects and 4 sessions.

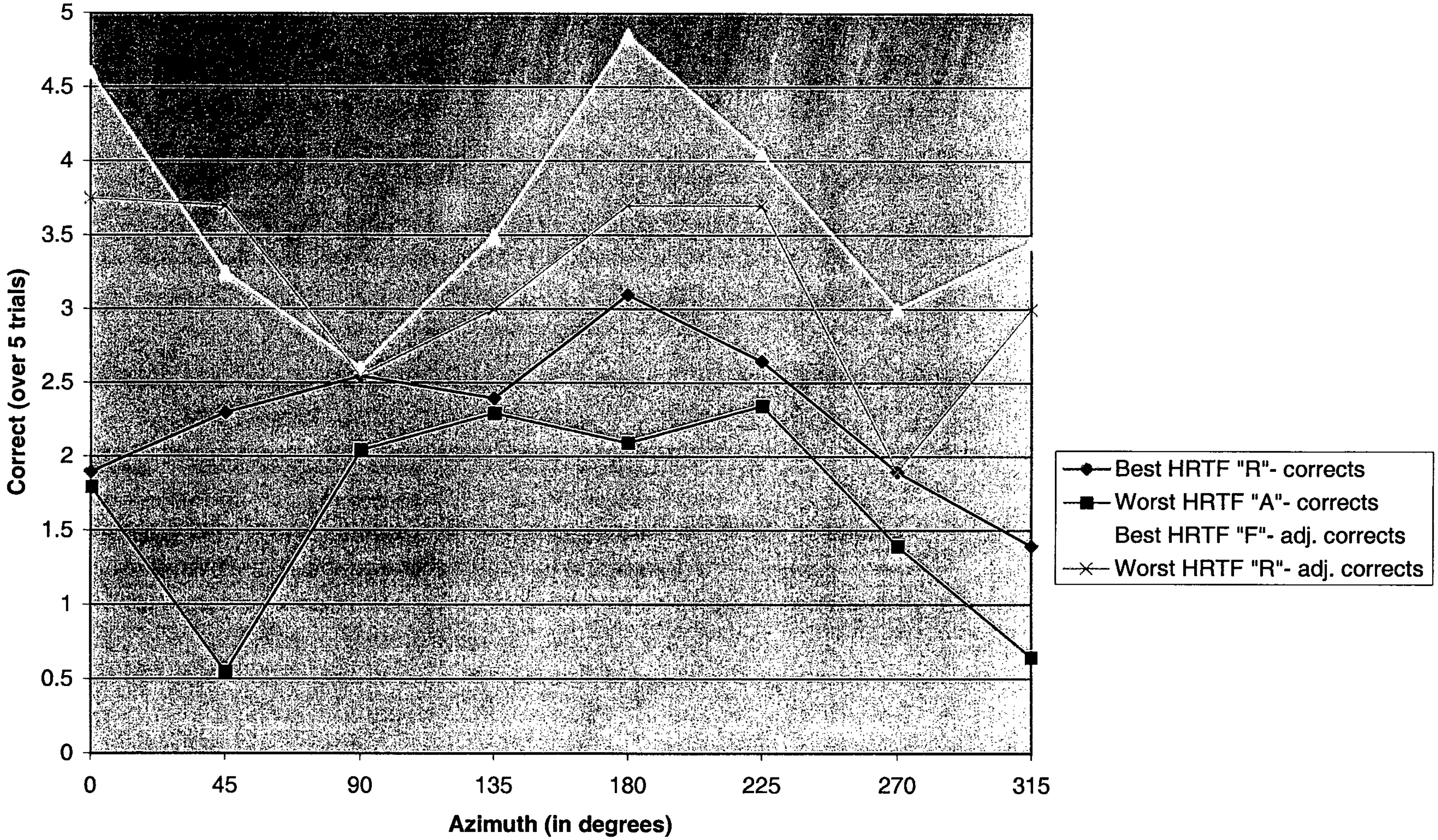

Figure 8b: A comparison of HRTFs that yielded best and worst localization accuracy for corrects and adjusted corrects in the high-pass 3 kHz condition of Experiment 1 (VAS - quiet). Differentiation is made between "corrects" and "adjusted corrects." Adjusted corrects resolve reversals by coding the subjects' responses as if they were made in the correct hemisphere. Data are averaged across 5 subjects and 4 sessions.

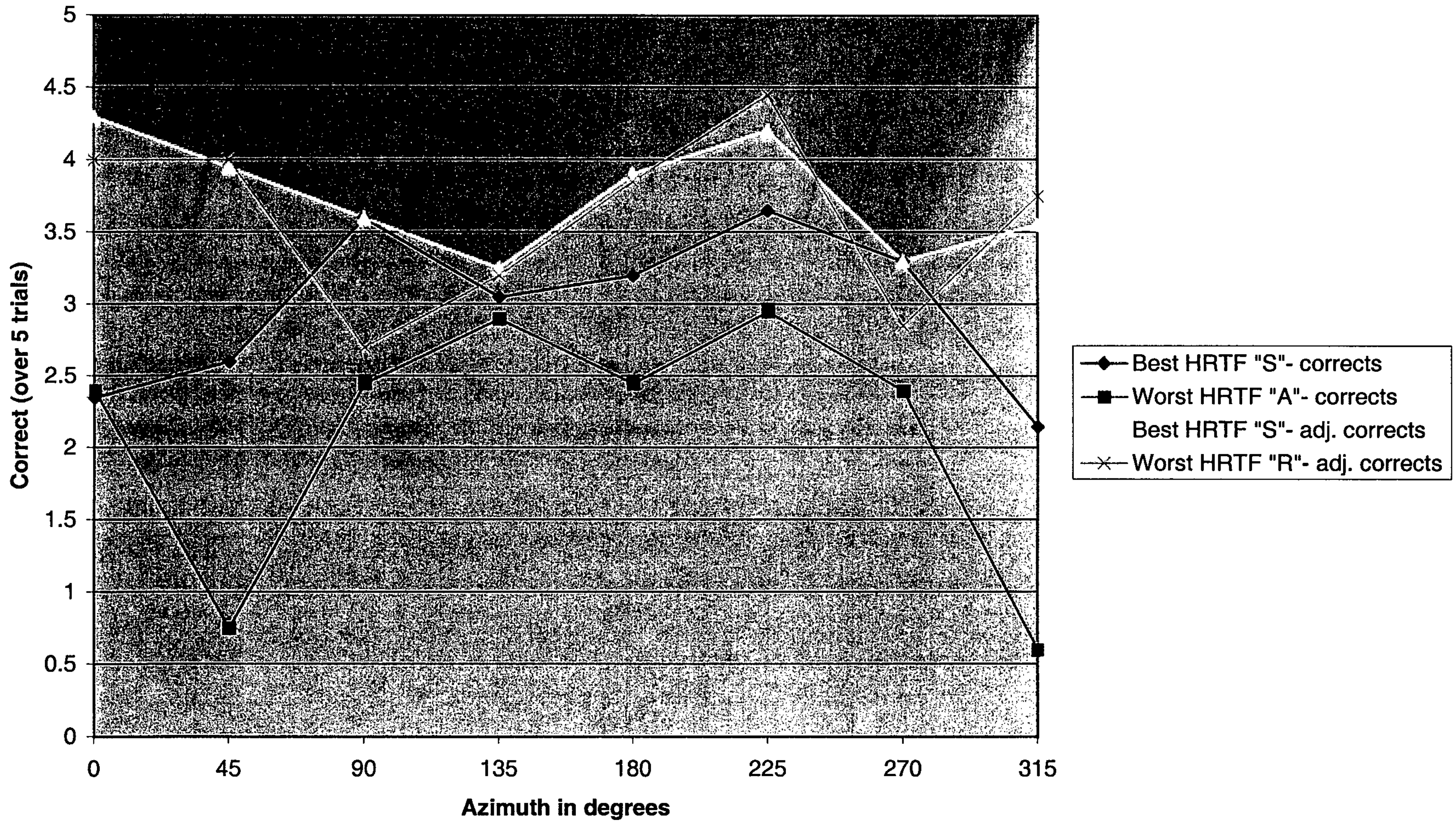

Figure 8c: A comparison of HRTFs that yielded best and worst localization accuracy for corrects and adjusted corrects in the broadband condition of Experiment 1 (VAS - quiet). Differentiation is made between "corrects" and "adjusted corrects." Adjusted corrects resolve reversals by coding the subjects' responses as if they were made in the correct hemisphere. Data are averaged across 5 subjects and 4 sessions.

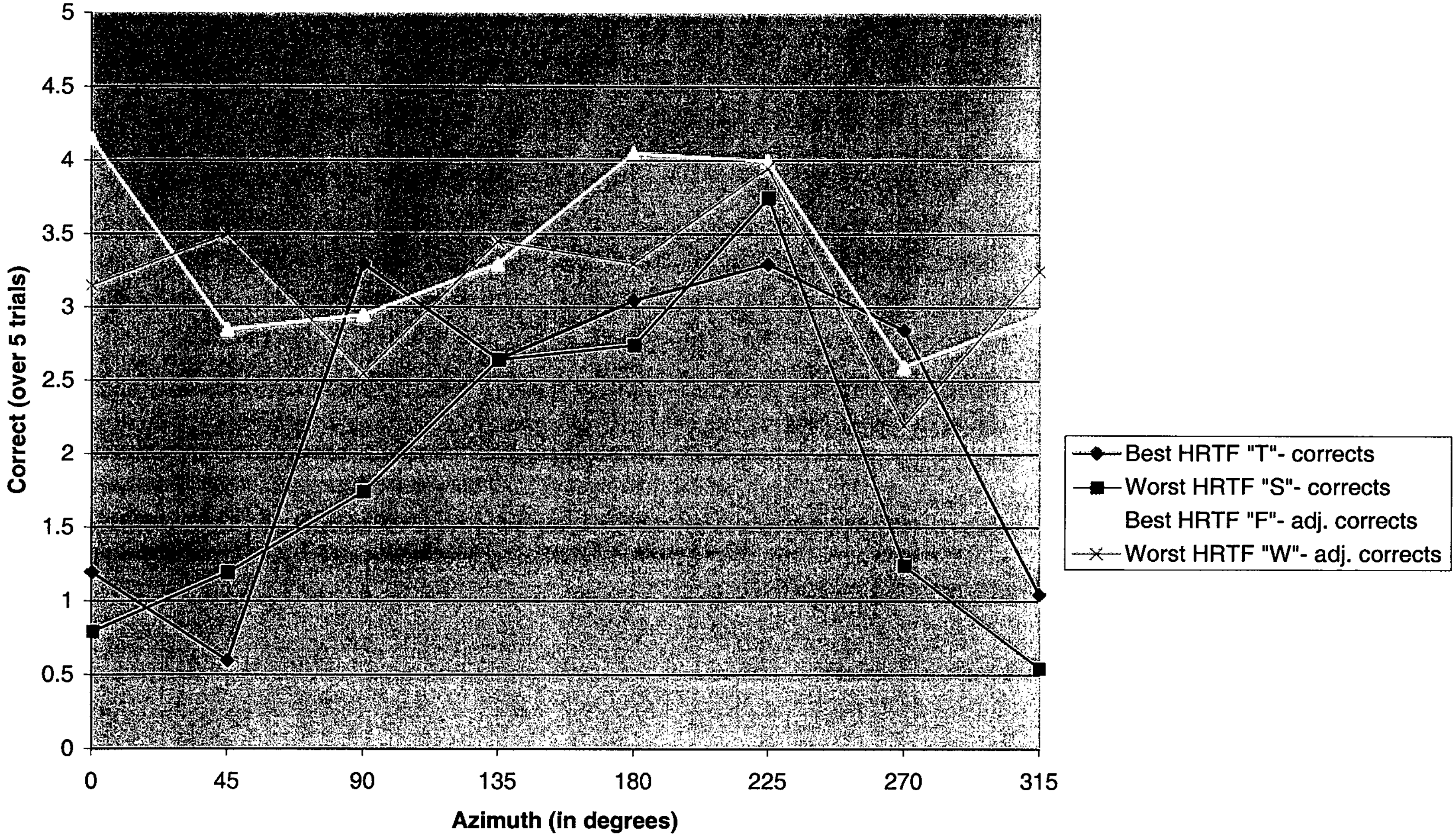

Figure 9a: A comparison of HRTFs that yielded best and worst localization accuracy for corrects and adjusted corrects in the low-pass 3 kHz condition of Experiment 2 (VAS - noise). Differentiation is made between "corrects" and "adjusted corrects." Adjusted corrects resolve reversals by coding the subjects' responses as if they were made in the correct hemisphere. Data are averaged across 5 subjects and 4 sessions.

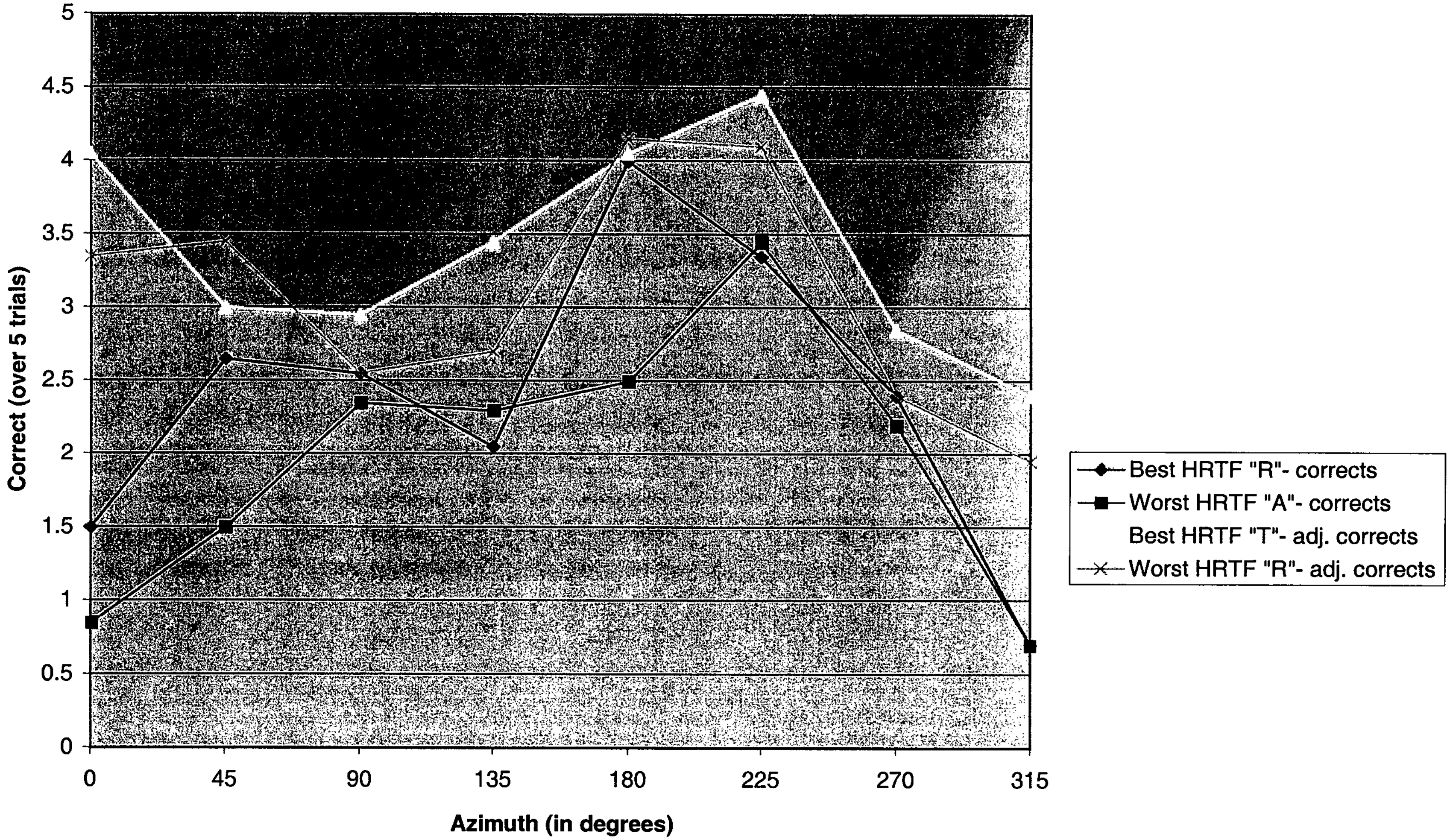

Figure 9b: A comparison of HRTFs that yielded best and worst localization accuracy for corrects and adjusted corrects in the high-pass 3 kHz condition of Experiment 2 (VAS - noise). Differentiation is made between "corrects" and "adjusted corrects." Adjusted corrects resolve reversals by coding the subjects' responses as if they were made in the correct hemisphere. Data are averaged across 5 subjects and 4 sessions.

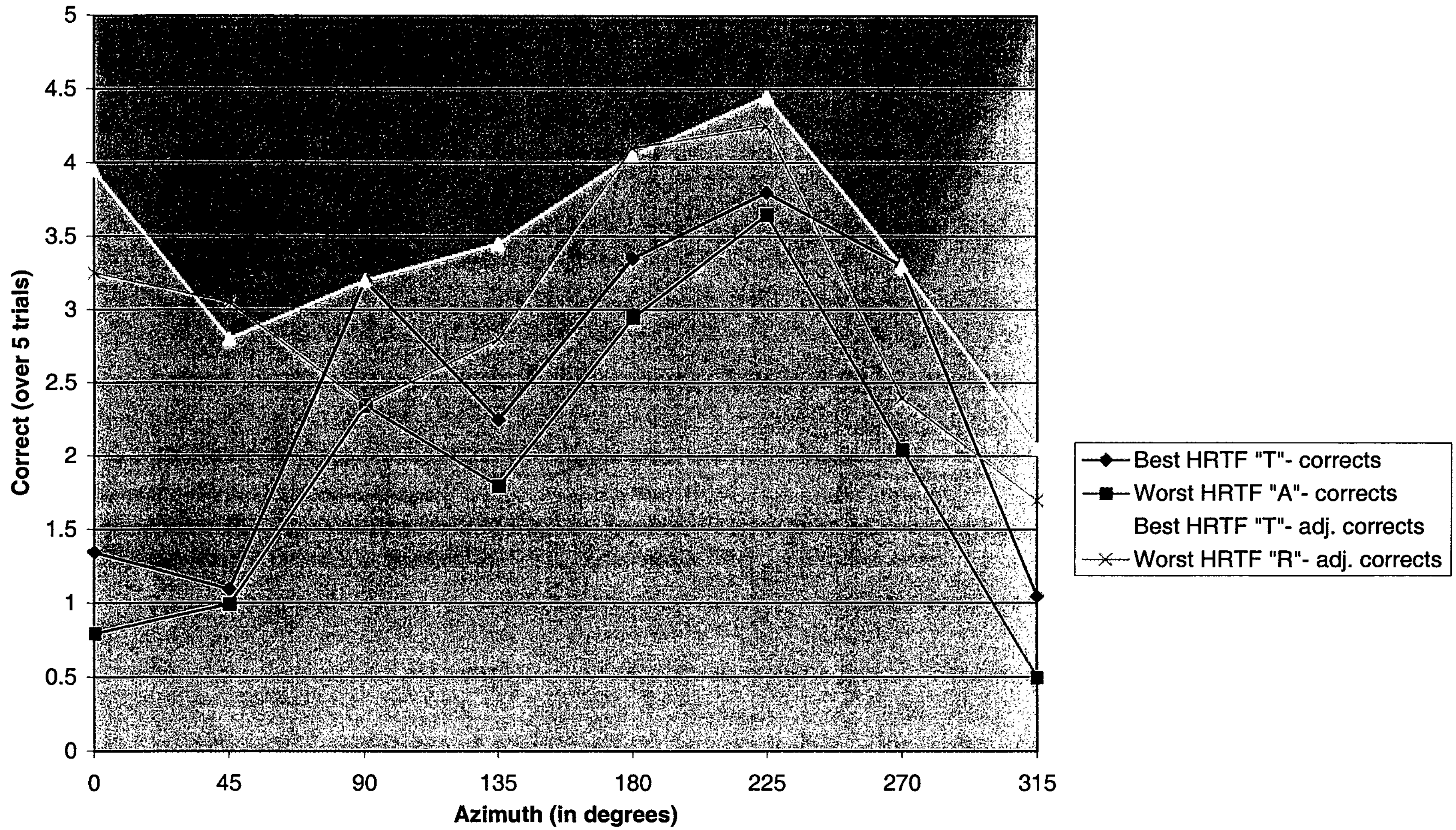

Figure 9c: A comparison of HRTFs that yielded best and worst localization accuracy for corrects and adjusted corrects in the broadband condition of Experiment 2 (VAS - noise). Differentiation is made between "corrects" and "adjusted corrects." Adjusted corrects resolve reversals by coding the subjects' responses as if they were made in the correct hemisphere. Data are averaged across 5 subjects and 4 sessions.

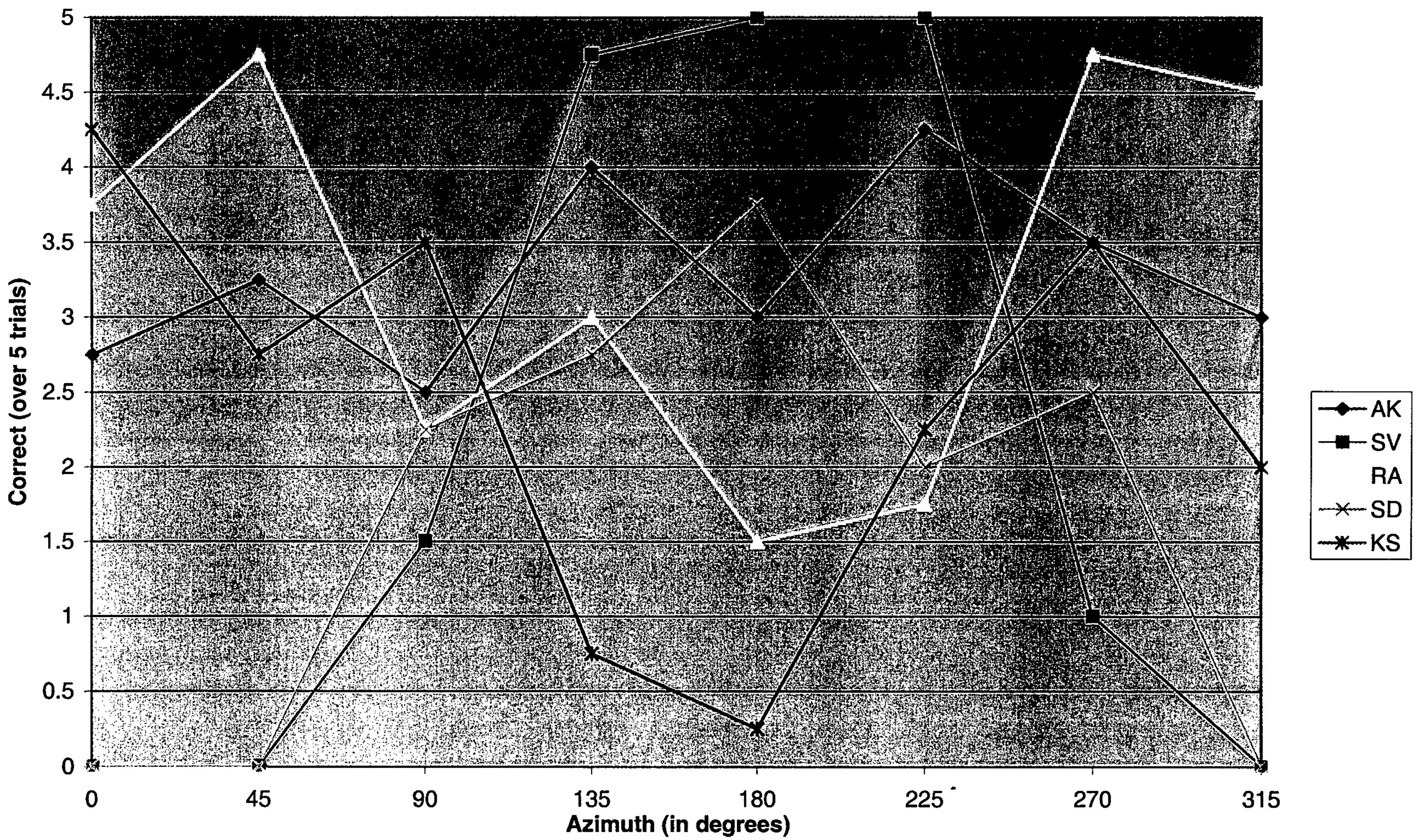

Figure 10a: A comparison of the variability of correct localization judgements among subjects in the low-pass 3 kHz condition of Experiment 1 (VAS - quiet). The comparison is based on HRTF "S" which yielded best localization accuracy in the low-pass 3 kHz condition for corrects (Figure 8a). Data are averaged across 4 sessions for each subject.

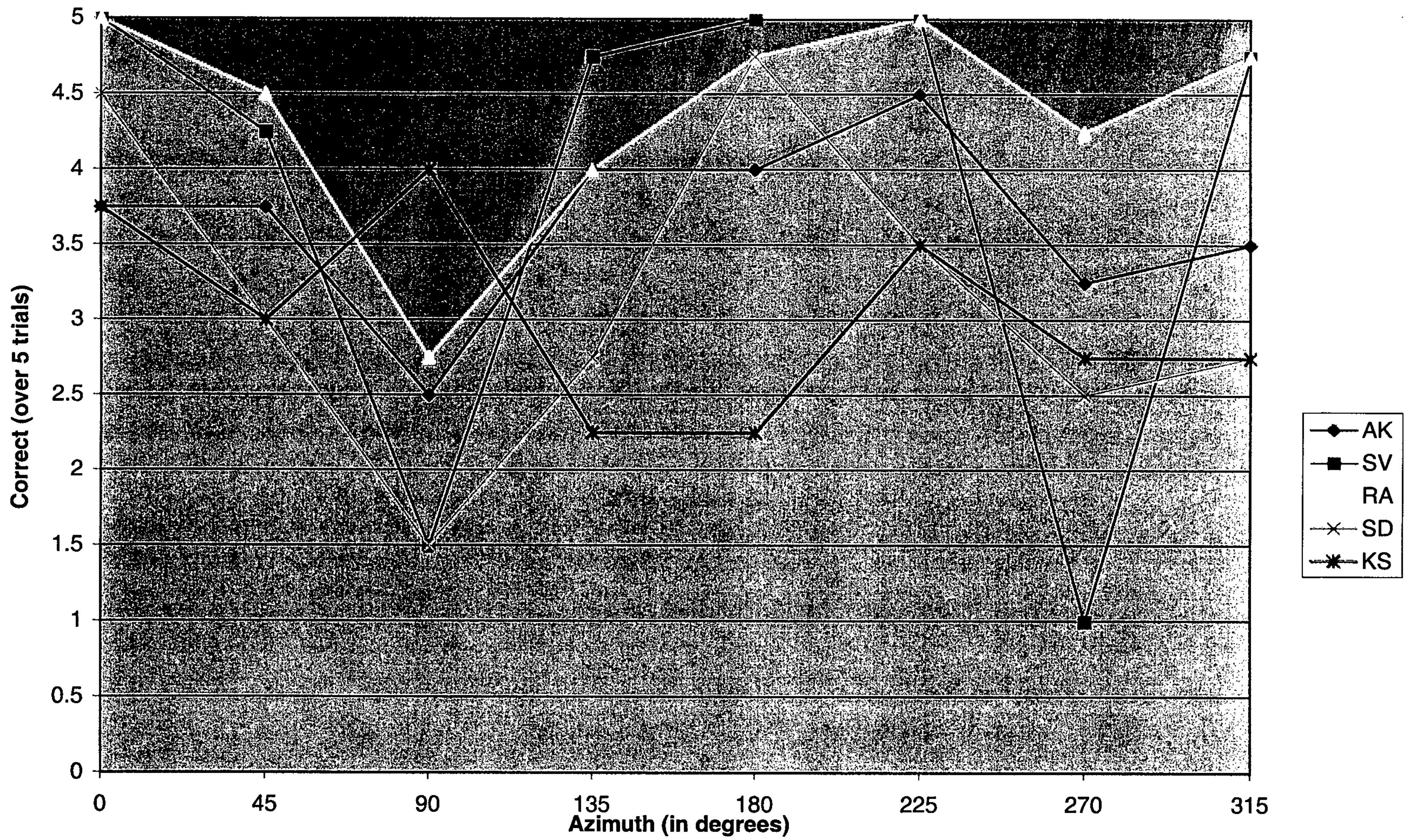

Figure 10b: A comparison of the variability of adjusted correct localization judgements among subjects in the low-pass 3 kHz condition of Experiment 1 (VAS - quiet). The comparison is based on HRTF "F" which yielded best localization accuracy in the low-pass 3 kHz condition for adjusted corrects (Figure 8a). Data are averaged across 4 sessions for each subject.

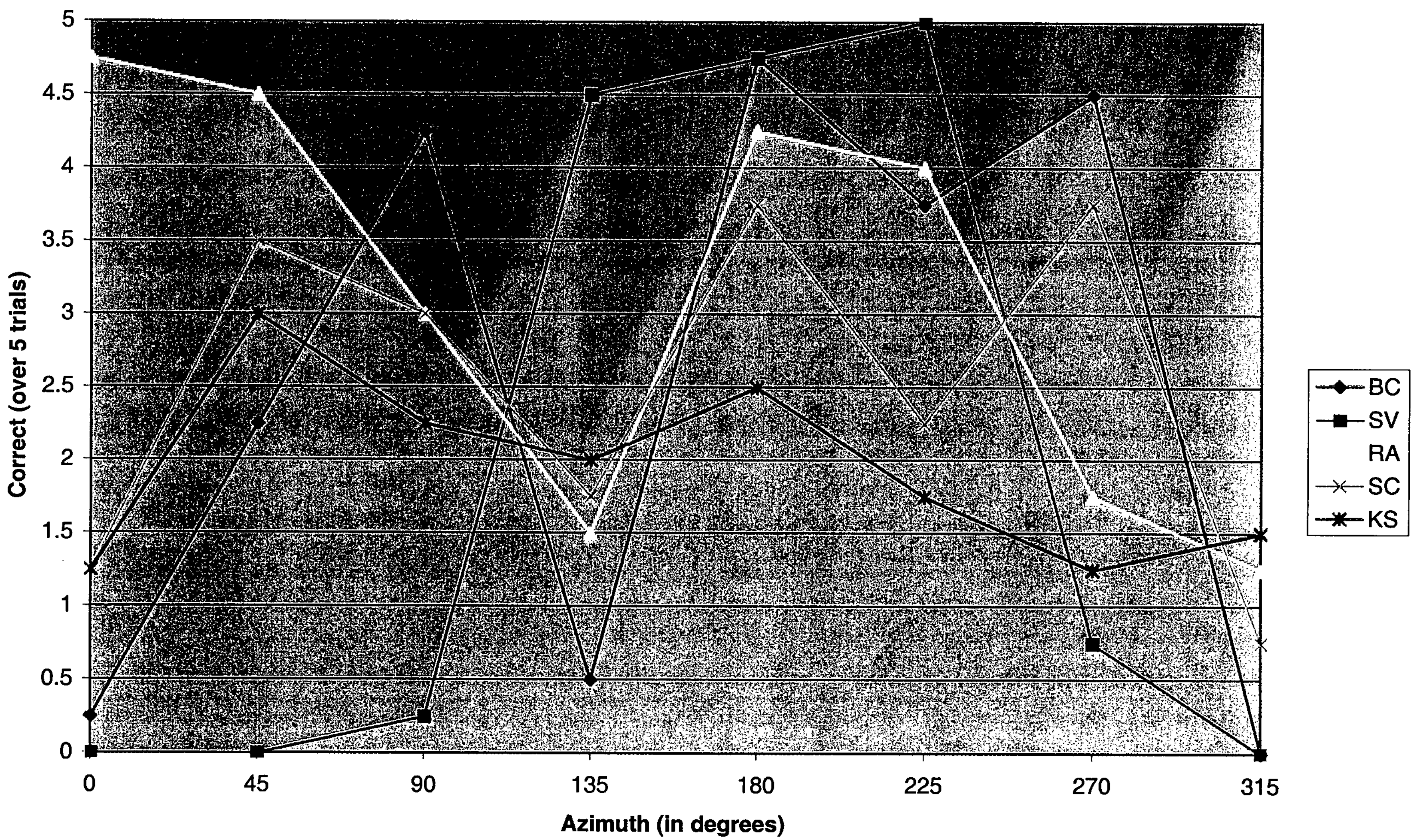

Figure 11a: A comparison of the variability of correct localization judgements among subjects in the high-pass 3 kHz condition of Experiment 2 (VAS - noise). The comparison is based on HRTF "R" which yielded best localization accuracy in the high-pass 3 kHz condition for corrects (Figure 9b). Data are averaged across 4 sessions for each subject.

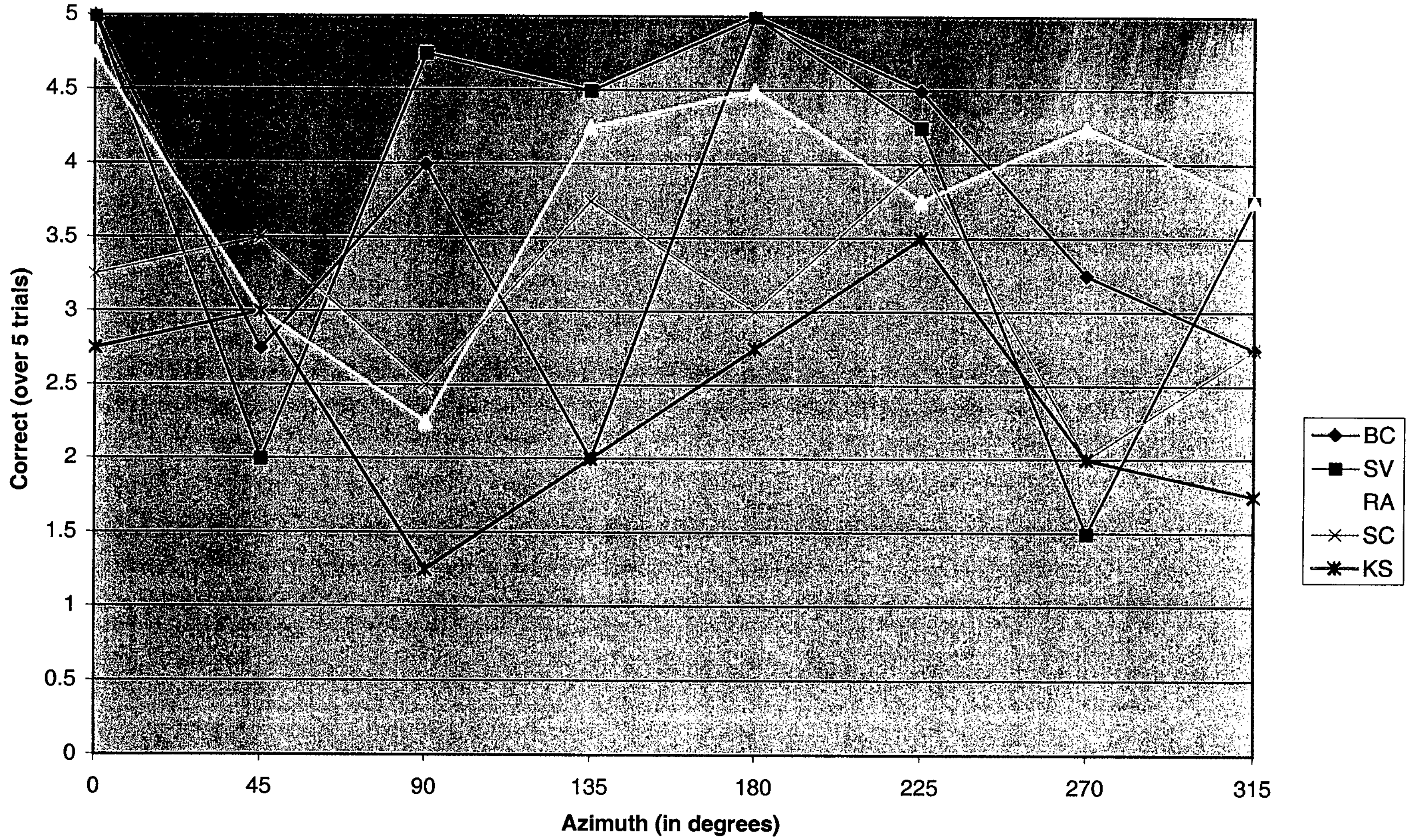

Figure 11b: A comparison of the variability of adjusted correct localization judgements among subjects in the low-pass 3 kHz condition of Experiment 2 (VAS - noise). The comparison is based on HRTF "F" which yielded best localization accuracy in the low-pass 3 kHz condition for adjusted corrects (Figure 9a). Data are averaged across 4 sessions for each subject.

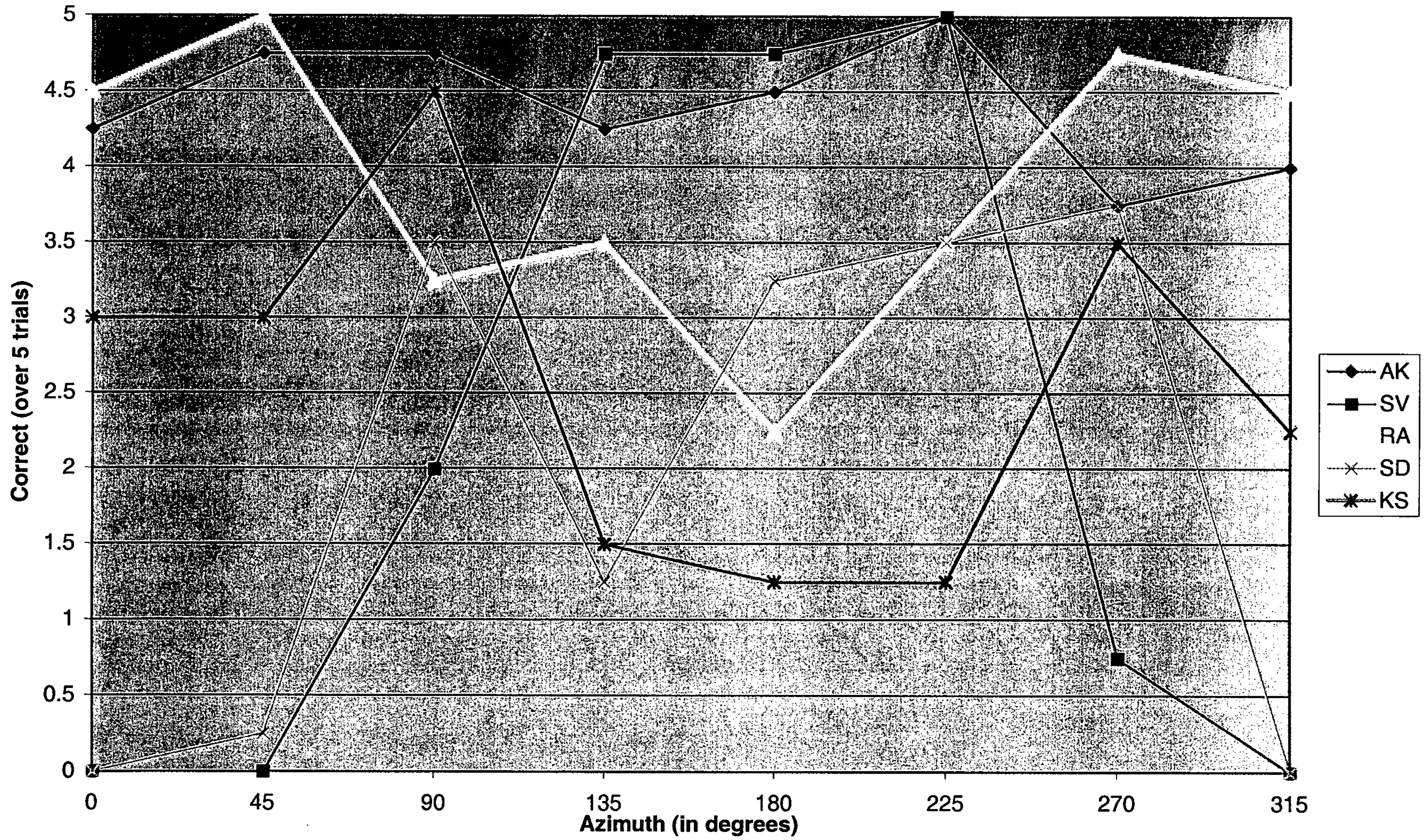

Figure 12a: A comparison of the variability of correct localization judgements among subjects in the broadband condition of Experiment 1 (VAS - quiet). The comparison is based on HRTF "S" which was measured on subject RA. Data are averaged across 4 sessions for each subject.

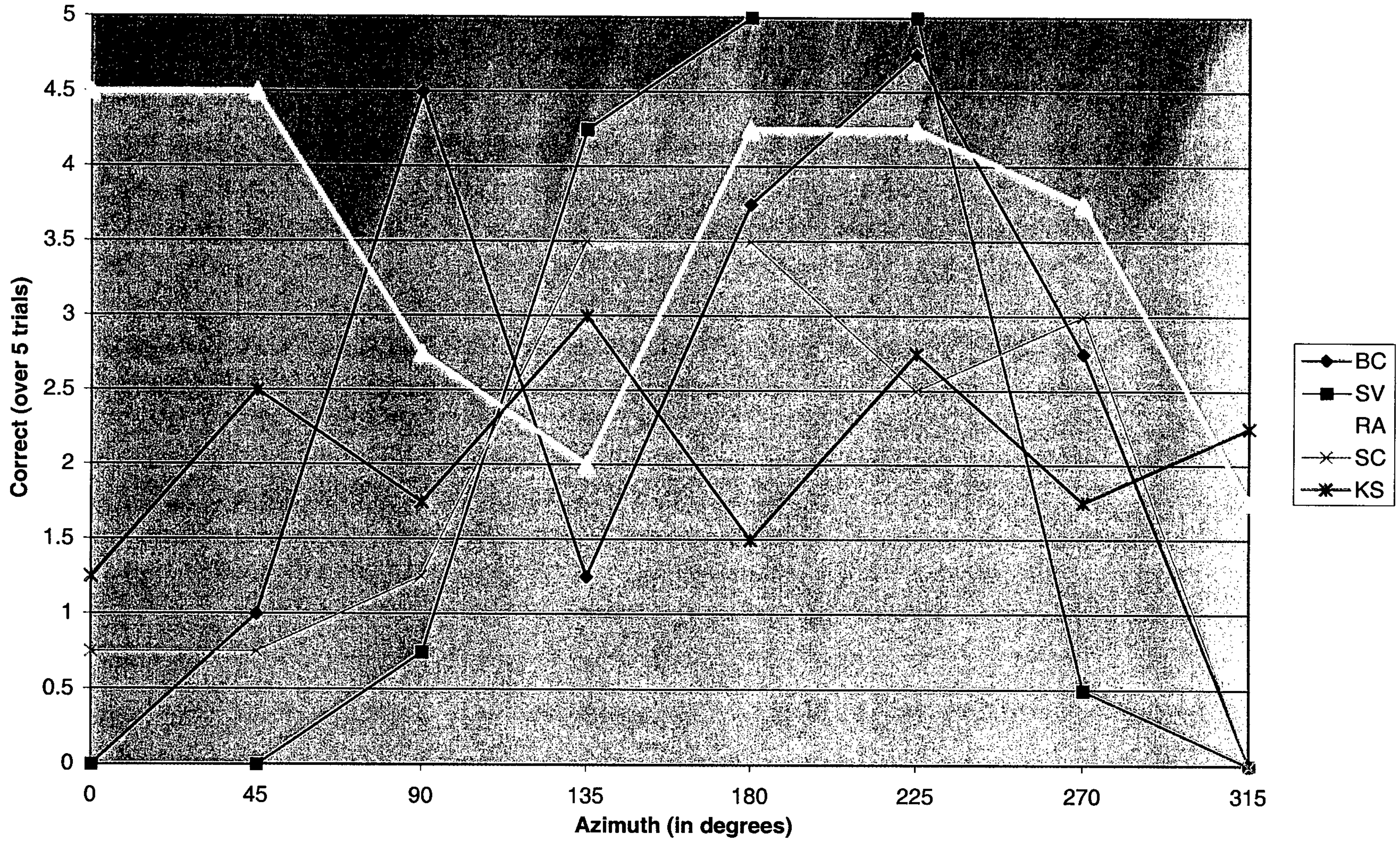

Figure 12b: A comparison of the variability of correct localization judgements among subjects in the broadband condition of Experiment 2 (VAS - noise). The comparison is based on HRTF "S" which was measured on subject RA. Data are averaged across 4 sessions for each subject.

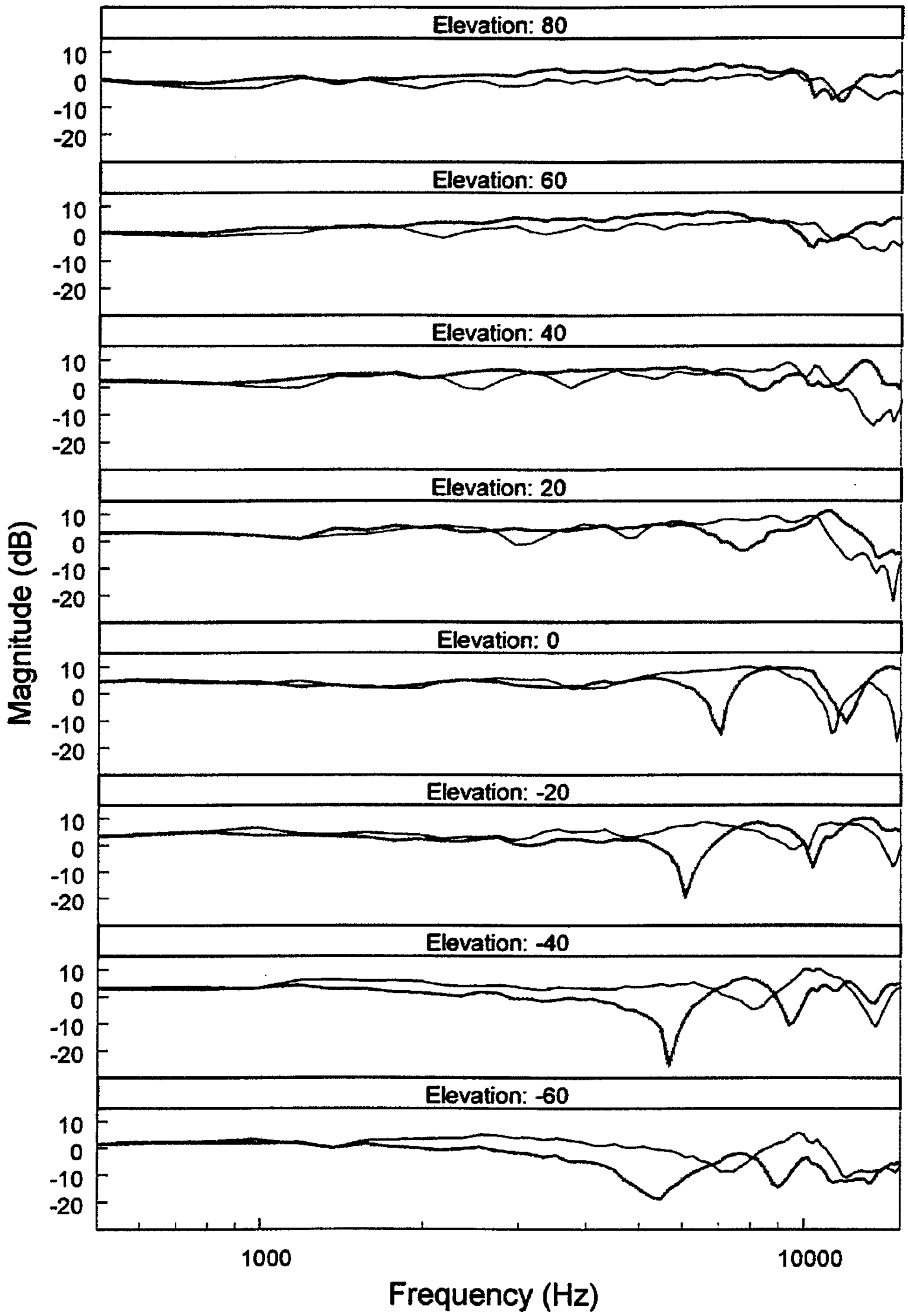

Figure 13: An example of HRTF variability. HRTFs were measured from one ear of two people, for a source direction at 90 degrees azimuth and elevations every 20 degrees from –60 to +80. Note that at some frequencies and elevations the two curves are as much as 30 dB apart. Courtesy of Fred Wightman.

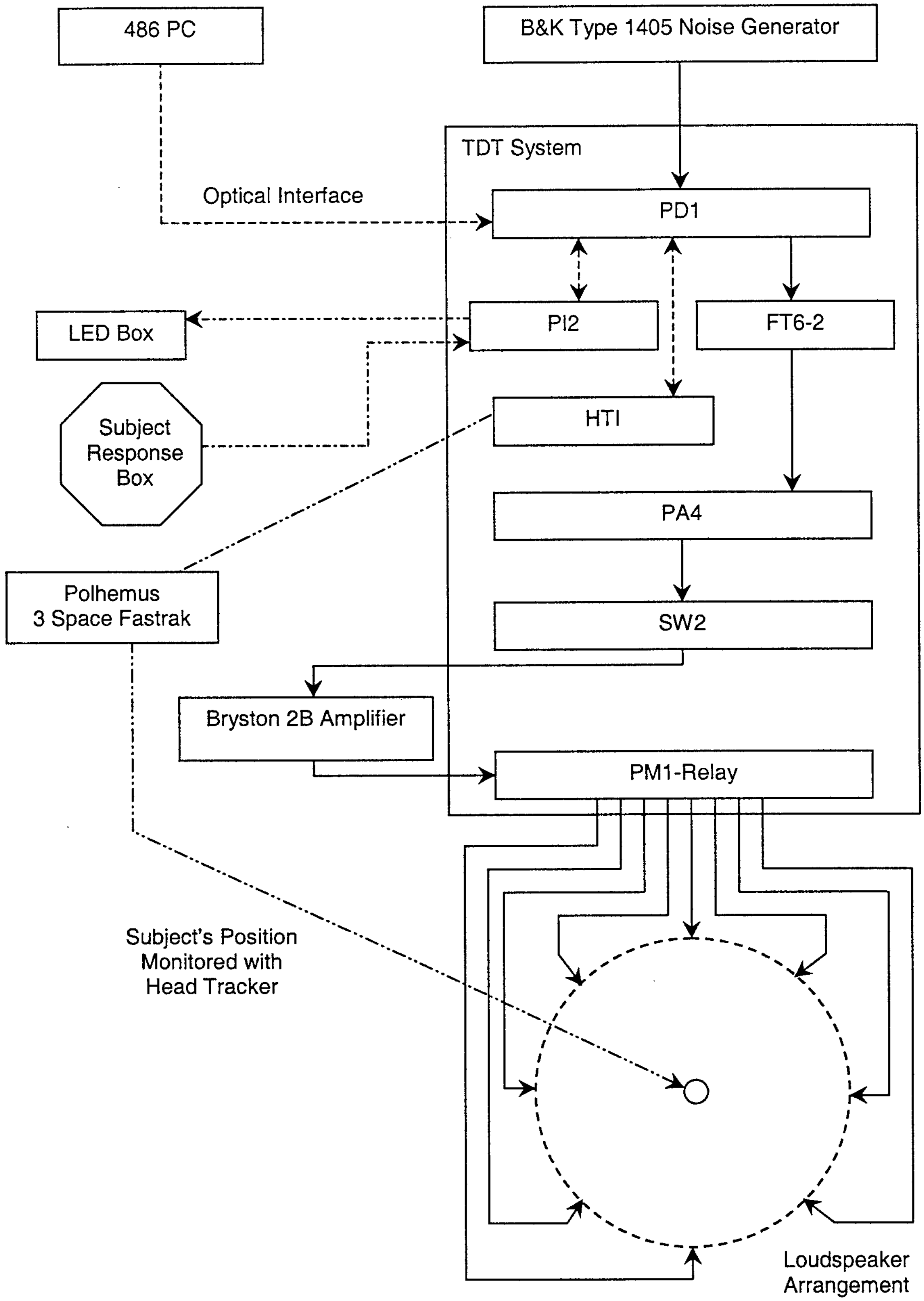

Figure 14: Hardware configuration for Experiments 3 and 4.

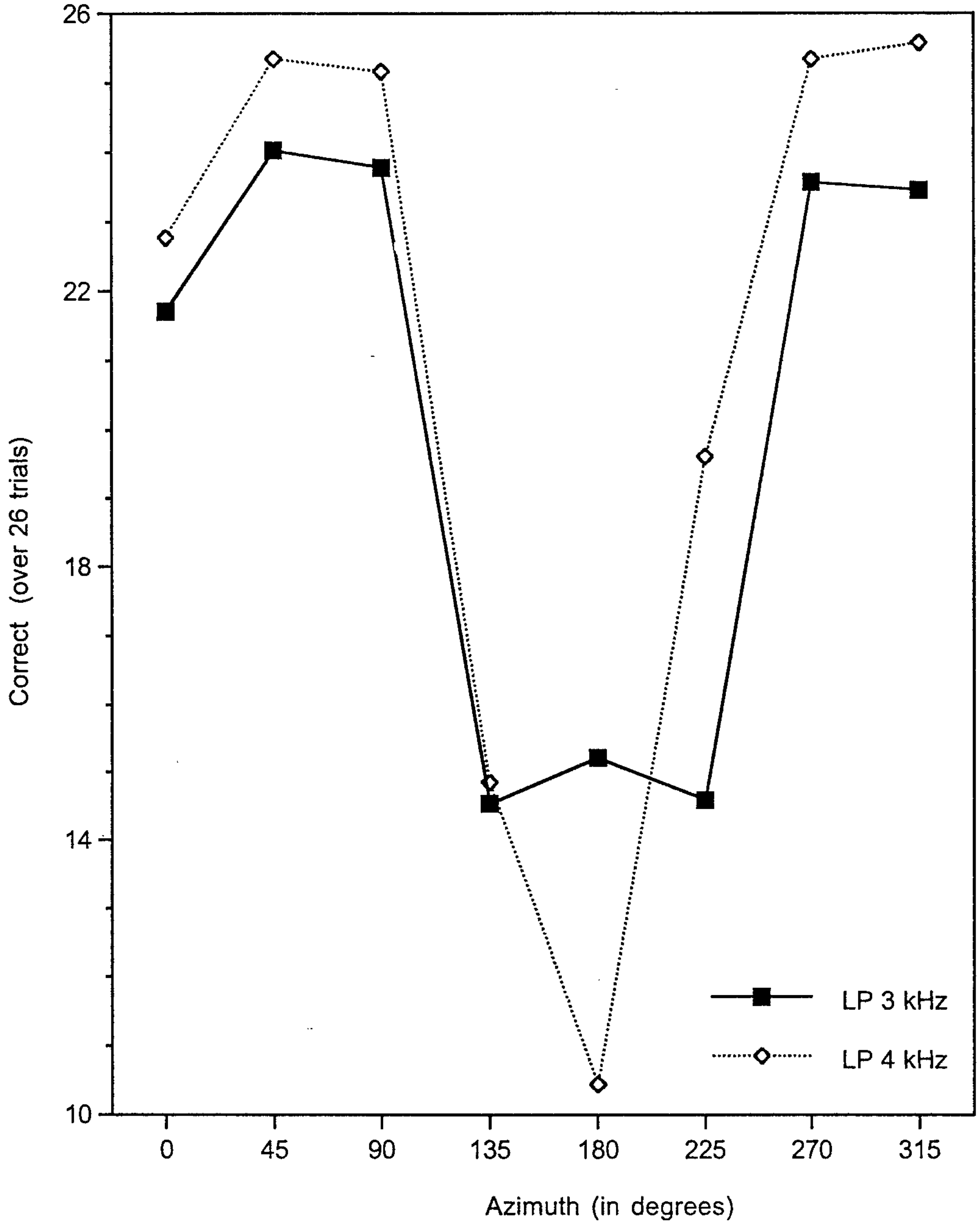

Figure 15a: Corrects. A comparison of localization accuracy with low-pass 3 kHz (Experiment 3) versus low-pass 4 kHz (Experiment 4). Data are averaged across 4 sessions for 17 subjects in the low-pass 3 kHz condition, and 13 subjects in the low-pass 4 kHz condition.

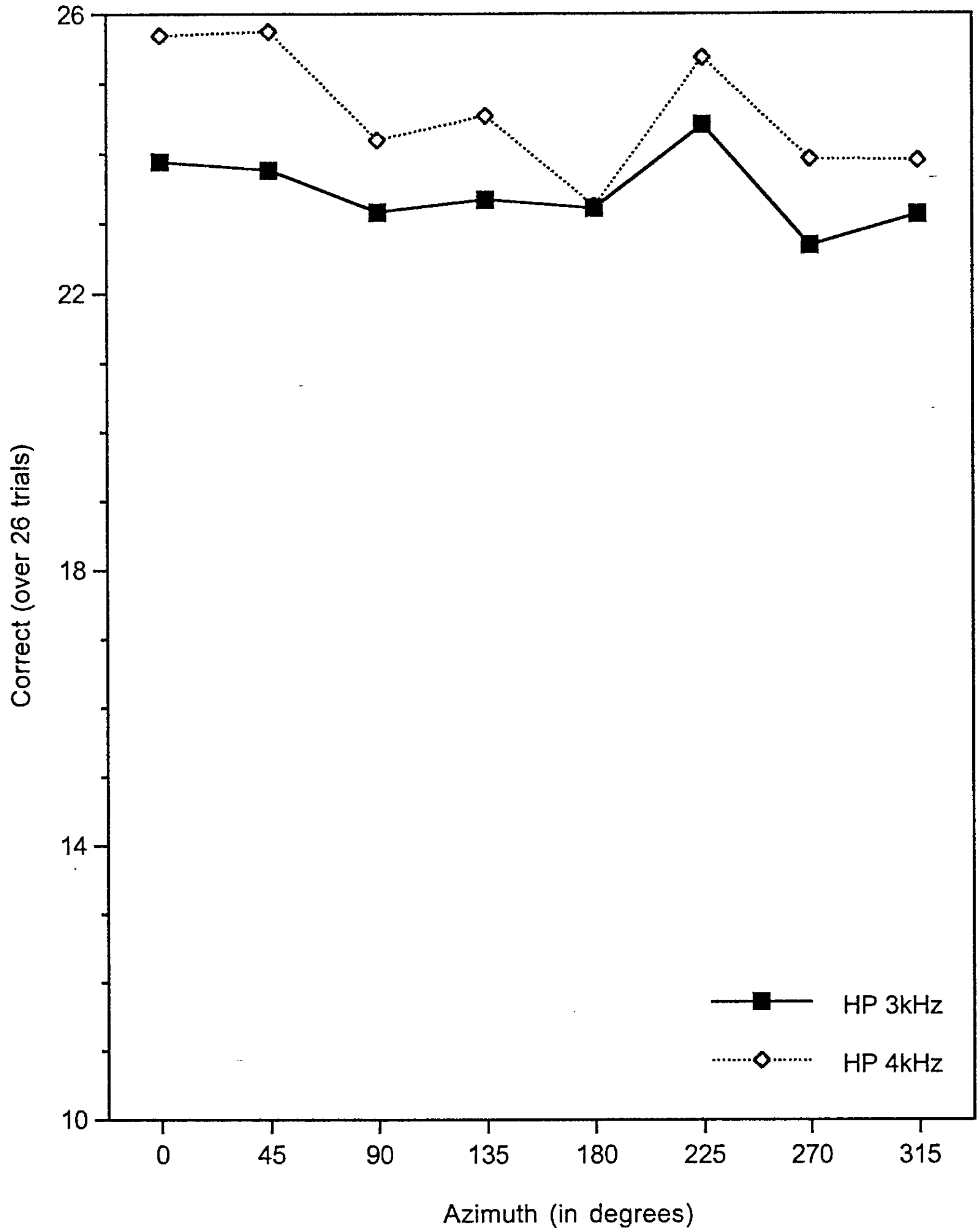

Figure 15b: Corrects. A comparison of localization accuracy with high-pass 3 kHz (Experiment 3) versus high-pass 4 kHz (Experiment 4). Data are averaged across 4 sessions for 17 subjects in the low-pass 3 kHz condition, and 13 subjects in the low-pass 4 kHz condition.

Table 1a: Localization performance in Experiment 1 (VAS - quiet). For each HRTF and stimulus condition the data are partitioned into corrects, adjusted corrects (i.e., sum of corrects and all reversals), errors, front-back and back-front reversals. Data are expressed as means and standard deviations (SD) and are averaged across 5 subjects and 4 sessions.

		HRTFs													
		A		F		K		R		S		T		W	
Stimulus	Category	Mean(%)	SD	Mean(%)	SD	Mean(%)	SD	Mean(%)	SD	Mean(%)	SD	Mean(%)	SD	Mean(%)	SD
BB	**Corrects**	**42.2**	**9.5**	**45.8**	**11.8**	**44.6**	**5.8**	**53.8**	**14.0**	**59.8**	**22.0**	**49.4**	**14.0**	**57.5**	**19.7**
	Adj. Corr.	73.8	12.9	74.8	12.1	74.5	11.0	72.0	9.6	75.1	13.9	75.3	13.4	74.8	11.3
	Errors	26.0	12.9	25.2	12.1	25.5	11.0	27.9	9.5	24.9	13.9	24.6	13.5	25.2	11.3
	Front-back	21.5	11.9	18.4	12.6	19.9	12.5	13.1	14.0	11.6	13.4	19.2	13.9	13.6	15.7
	Back-front	10.0	9.6	10.6	10.0	10.0	11.3	5.1	8.3	3.6	5.0	6.6	8.9	3.6	5.0
HP 3 kHz	**Corrects**	**33.0**	**5.4**	**45.6**	**10.2**	**36.2**	**5.6**	**45.5**	**14.5**	**45.6**	**13.1**	**39.9**	**7.8**	**42.1**	**14.9**
	Adj. Corr.	66.2	9.4	73.2	10.6	66.0	9.6	63.3	13.8	66.0	10.3	69.6	12.1	66.9	13.5
	Errors	33.8	9.5	26.8	10.5	34.0	9.6	36.8	13.8	34.0	10.3	30.4	12.1	33.0	13.5
	Front-back	19.8	12.8	16.2	13.3	17.1	14.6	12.1	13.9	15.2	15.4	19.9	13.2	16.1	14.0
	Back-front	13.5	14.6	11.4	10.7	12.0	12.2	5.6	7.1	5.1	7.9	9.8	11.1	8.5	10.8
LP 3 kHZ	**Corrects**	**45.5**	**9.1**	**47.9**	**9.3**	**45.9**	**12.4**	**50.7**	**14.1**	**51.1**	**14.3**	**45.0**	**10.9**	**47.4**	**11.4**
	Adj. Corr.	71.5	11.7	72.5	11.0	72.8	12.9	72.0	8.8	70.6	15.7	69.5	15.6	71.6	12.2
	Errors	28.5	11.7	27.5	11.0	27.0	12.7	28.0	8.7	29.4	15.7	30.4	15.5	28.4	12.2
	Front-back	20.4	13.0	18.2	11.8	18.1	13.2	15.6	14.8	13.5	13.1	19.2	12.1	16.4	14.1
	Back-front	5.5	6.6	6.2	7.7	8.9	8.2	5.6	6.4	6.0	7.9	5.1	6.2	7.9	9.1

Table 1b: Localization performance in Experiment 2 (VAS - noise). For each HRTF and stimulus condition the data are partitioned into corrects, adjusted corrects (i.e., sum of corrects and all reversals), errors, front-back and back-front reversals. Data are expressed as means and standard deviations (SD) and are averaged across 5 subjects and 4 sessions.

		HRTFs													
		A		F		K		R		S		T		W	
Stimulus	Category	Mean(%)	SD	Mean(%)	SD	Mean(%)	SD	Mean(%)	SD	Mean(%)	SD	Mean(%)	SD	Mean(%)	SD
BB	**Corrects**	**37.8**	**6.8**	**43.6**	**14.4**	**42.4**	**8.2**	**46.5**	**10.5**	**46.6**	**13.0**	**48.5**	**7.5**	**40.2**	**10.2**
	Adj. Corr.	60.7	10.0	64.9	9.7	66.3	13.2	59.7	7.5	62.4	9.2	68.1	7.3	63.6	13.1
	Errors	39.2	10.0	35.1	9.7	33.6	13.1	40.2	7.5	37.6	9.2	31.9	7.5	36.4	13.1
	Front-back	17.8	9.8	13.8	7.2	17.9	12.4	10.1	8.7	11.9	10.3	13.2	9.0	18.4	10.9
	Back-front	5.2	5.2	7.5	6.5	5.9	6.3	3.0	2.8	3.5	3.5	6.4	5.6	5.0	5.1
HP 3 kHz	**Corrects**	**39.6**	**4.6**	**42.5**	**7.8**	**40.9**	**7.2**	**48.0**	**10.0**	**46.4**	**13.9**	**47.0**	**9.8**	**44.5**	**10.7**
	Adj. Corr.	61.2	7.6	64.5	4.5	63.5	9.8	61.6	5.6	61.5	11.1	68.0	8.7	65.3	11.3
	Errors	38.6	7.4	35.5	4.5	36.5	9.8	38.4	5.6	38.5	11.0	32.0	8.7	34.6	11.3
	Front-back	16.2	9.5	14.1	9.3	16.1	11.3	9.8	10.2	11.1	10.3	14.5	8.6	14.5	7.3
	Back-front	5.2	5.6	7.8	6.0	6.5	6.9	3.9	4.6	4.0	3.5	6.4	5.9	6.4	6.4
LP 3 kHz	**Corrects**	**43.4**	**7.2**	**41.4**	**8.4**	**41.5**	**9.0**	**40.0**	**14.0**	**36.8**	**8.5**	**45.0**	**9.8**	**38.5**	**8.9**
	Adj. Corr.	63.7	11.7	67.1	12.5	63.1	11.2	62.4	15.1	57.0	13.2	66.1	14.2	63.4	12.8
	Errors	36.2	11.6	32.8	12.3	36.8	10.9	37.5	15.1	43.0	13.3	33.8	14.1	36.6	12.8
	Front-back	16.5	12.5	16.6	9.2	17.2	12.3	16.4	12.6	15.8	11.8	16.1	10.8	18.5	12.4
	Back-front	3.5	3.2	9.0	9.4	4.1	4.0	5.5	5.1	3.4	3.6	4.9	5.2	5.8	5.8

Table 1c: Localization performance for subjects KS, RA, and SV who participated in both Experiment 1 (VAS - quiet) and Experiment 2 (VAS - noise). For each HRTF and stimulus condition, the data are partitioned into corrects, errors and sum of reversals. Means are expressed as percentages and are averaged across 4 sessions for each subject.

			Stimulus								
			BB			HP 3 kHz			LP 3 kHz		
			Subject								
Experiment	HRTFs	Data	KS	RA	SV	KS	RA	SV	KS	RA	SV
1	A	**Corrects**	**38.8**	**56.2**	**42.5**	**34.4**	**23.8**	**36.9**	**41.9**	**52.5**	**40.0**
		Errors	40.6	7.5	23.1	48.8	31.9	26.9	46.9	16.9	23.8
		Sum of Reversals	20.6	36.2	34.4	16.9	44.3	36.2	11.2	30.6	36.2
	F	**Corrects**	**33.8**	**61.9**	**40.0**	**36.2**	**62.5**	**43.8**	**43.8**	**60.0**	**43.1**
		Errors	39.4	11.2	23.8	35.6	12.5	20.0	39.4	12.5	21.9
		Sum of Reversals	26.9	26.8	36.2	28.1	25.0	36.2	16.9	27.5	35.0
	K	**Corrects**	**39.4**	**48.1**	**44.4**	**36.9**	**37.5**	**38.1**	**41.2**	**63.7**	**41.2**
		Errors	40.6	12.5	18.8	44.4	25.0	24.4	41.2	11.2	23.8
		Sum of Reversals	20.0	39.4	36.9	18.7	37.5	37.5	17.5	25.0	35.0
	R	**Corrects**	**50.6**	**60.6**	**46.2**	**35.6**	**51.9**	**37.5**	**48.8**	**61.3**	**37.5**
		Errors	38.8	18.8	21.9	52.5	27.5	26.9	38.8	18.8	26.9
		Sum of Reversals	10.6	20.6	31.9	11.9	20.6	35.6	12.4	20.0	35.6
	S	**Corrects**	**50.6**	**78.1**	**43.1**	**45.0**	**52.5**	**38.8**	**48.1**	**65.6**	**43.1**
		Errors	35.6	11.2	25.6	46.9	25.0	25.6	43.8	10.6	23.1
		Sum of Reversals	13.8	10.6	31.2	8.1	22.6	35.6	8.1	23.8	33.8
	T	**Corrects**	**45.6**	**64.4**	**38.8**	**33.1**	**45.6**	**35.6**	**38.8**	**55.6**	**41.9**
		Errors	40.0	7.5	24.4	47.5	14.4	27.5	51.2	16.2	21.2
		Sum of Reversals	14.4	28.1	36.9	19.3	40.0	36.9	10.0	28.1	36.9
	W	**Corrects**	**51.9**	**74.4**	**41.9**	**33.1**	**42.5**	**36.9**	**43.1**	**61.9**	**41.9**
		Errors	38.8	12.5	25.6	49.4	19.4	25.6	44.4	15.0	23.8
		Sum of Reversals	9.4	13.1	32.5	17.5	38.1	36.9	12.5	23.1	34.4
2	A	**Corrects**	**30.6**	**43.8**	**33.8**	**34.4**	**45.6**	**36.2**	**33.1**	**53.1**	**43.8**
		Errors	53.8	33.8	35.0	47.5	30.0	31.9	50.6	33.1	22.5
		Sum of Reversals	15.6	22.5	31.2	17.5	24.4	31.9	16.2	13.7	33.8
	F	**Corrects**	**26.2**	**62.5**	**47.5**	**38.1**	**55.0**	**38.8**	**31.2**	**45.0**	**50.0**
		Errors	46.2	23.8	30.6	39.4	30.0	32.5	51.9	23.8	23.1
		Sum of Reversals	27.5	13.7	21.9	22.5	15.0	28.7	16.2	31.2	26.9
	K	**Corrects**	**36.9**	**37.5**	**43.8**	**31.9**	**38.8**	**40.6**	**28.7**	**53.1**	**40.6**
		Errors	45.0	40.0	20.0	45.6	33.1	24.4	51.9	35.6	26.9
		Sum of Reversals	18.1	22.5	36.2	22.6	28.1	35.0	18.7	11.3	32.5
	R	**Corrects**	**33.1**	**61.3**	**41.2**	**38.8**	**62.5**	**38.1**	**22.5**	**60.6**	**44.4**
		Errors	51.9	35.0	36.2	46.9	33.8	35.0	56.9	22.5	24.4
		Sum of Reversals	15.0	3.7	22.5	14.4	3.7	26.9	20.5	16.9	31.2
	S	**Corrects**	**41.9**	**69.4**	**38.8**	**26.2**	**65.0**	**43.8**	**25.6**	**46.2**	**36.2**
		Errors	41.2	25.0	34.4	55.6	28.1	29.4	58.8	37.5	32.5
		Sum of Reversals	16.9	5.7	26.9	18.1	6.9	26.9	15.7	16.2	31.2
	T	**Corrects**	**38.8**	**58.1**	**44.4**	**32.5**	**54.4**	**41.2**	**34.4**	**60.6**	**45.0**
		Errors	43.8	24.4	28.1	46.9	24.4	30.0	52.5	18.1	23.1
		Sum of Reversals	17.6	17.4	27.5	20.6	21.2	28.7	13.2	21.3	31.9
	W	**Corrects**	**28.1**	**55.6**	**36.9**	**35.0**	**61.9**	**36.9**	**26.9**	**50.6**	**39.4**
		Errors	53.8	22.5	30.0	48.1	16.9	37.5	53.1	26.9	25.0
		Sum of Reversals	18.1	21.9	33.1	16.8	21.3	25.6	20.0	22.5	35.6

Table 2: Localization performance for Experiment 3 (free-field - 3 kHz cutoff frequency) and Experiment 4 (free-field - 4 kHz cutoff frequency). For each stimulus condition the data are partitioned into corrects, adjusted corrects (i.e., sum of corrects and all reversals), errors, front-back and back-front reversals. Data are expressed as means and standard deviations (SD) and are averaged across 4 sessions for each of 17 subjects (Experiment 3) and 13 subjects (Experiment 4).

		Stimulus					
		BB		HP		LP	
Experiment	Category	Mean (%)	SD	Mean (%)	SD	Mean (%)	SD
3 - 3kHz	**Corrects**	**93.4**	**15.2**	**90.2**	**16.5**	**77.3**	**17.3**
	Adj. Corr.	95.9	7.4	93.5	9.1	94.5	8.4
	Errors	3.5	7.3	5.8	8.6	5.1	8.1
	Front-back	2.2	8.4	2.2	8.1	3.3	7.6
	Back-front	0.2	0.6	1.1	2.4	13.8	12.8
4 - 4kHz	**Corrects**	**98.1**	**2.4**	**94.5**	**4.9**	**85.6**	**12.5**
	Adj. Corr.	98.4	2.4	95.9	3.7	97.9	2.3
	Errors	1.5	2.3	3.8	3.6	2.0	2.3
	Front-back	0.1	0.2	0.1	0.3	1.7	1.9
	Back-front	0.2	0.3	1.3	2.7	10.5	12.7

Table 3a: A comparison of the variability in localization accuracy among subjects in Experiment 3 (3 kHz cutoff frequency). For each stimulus condition the data are partitioned into corrects, adjusted corrects (i.e., sum of corrects and all reversals), errors, front-back reversals and back-front reversals. Data are expressed as means and are averaged across 4 sessions for each subject.

Subject	Stimulus														
	BB					HP 3 kHz					LP 3 kHz				
	Corrects	Adj Corr	Errors	F-B	B-F	**Corrects**	Adj Corr	Errors	F-B	B-F	**Corrects**	Adj Corr	Errors	F-B	B-F
AA*	**96.4**	99.2	0.5	0.4	2.4	**88.6**	98.2	0.7	0.1	9.5	**67.7**	99.3	0.7	0.6	30.9
AD	**96.6**	96.6	1.6	0.0	0.0	**92.4**	93.1	3.4	0.1	0.6	**69.1**	88.7	10.9	0.0	19.6
AS	**98.6**	98.7	0.6	0.1	0.0	**99.0**	99.1	0.7	0.0	0.1	**61.5**	98.5	0.5	0.0	37.0
BC*	**99.8**	99.9	0.0	0.1	0.0	**99.3**	99.3	0.7	0.0	0.0	**85.9**	99.6	0.4	0.2	13.5
BK	**99.5**	99.5	0.0	0.0	0.0	**98.4**	98.4	1.4	0.0	0.0	**73.7**	99.6	0.4	10.5	15.4
CT	**99.8**	99.8	0.0	0.0	0.0	**99.4**	99.6	0.4	0.0	0.2	**88.0**	99.5	0.4	1.3	10.2
DH*	**99.3**	99.3	0.0	0.0	0.0	**98.1**	98.9	1.1	0.7	0.0	**98.4**	99.9	0.0	0.5	1.0
EP*	**98.6**	98.7	1.2	0.0	0.0	**95.1**	95.1	4.9	0.0	0.0	**63.1**	98.8	0.7	0.0	35.7
JM*	**96.8**	96.8	0.8	0.0	0.0	**92.5**	94.4	4.7	0.0	1.9	**77.3**	93.4	5.4	0.0	16.0
KD	**82.3**	84.2	13.9	1.0	0.4	**73.1**	76.5	21.3	0.2	2.5	**55.5**	86.3	12.9	0.4	29.1
KS	**99.4**	99.5	0.5	0.0	0.0	**98.8**	98.8	1.0	0.0	0.0	**91.7**	98.5	1.3	1.3	5.5
RA*	**93.5**	94.9	4.9	0.4	1.0	**85.6**	91.0	9.0	2.2	3.2	**89.3**	94.9	5.0	1.8	3.7
RT*	**98.7**	98.7	1.0	0.0	0.0	**96.0**	96.0	4.0	0.0	0.0	**76.8**	94.7	5.0	3.0	14.9
SC	**97.4**	97.6	2.3	0.0	0.2	**95.8**	96.4	3.0	0.6	0.0	**93.3**	96.0	4.0	0.8	1.9
SD	**36.5**	71.1	28.8	34.6	0.0	**31.7**	65.2	32.8	33.5	0.0	**34.4**	65.7	32.9	31.2	0.0
SK	**98.6**	98.6	1.4	0.0	0.0	**96.4**	96.4	3.1	0.0	0.0	**94.2**	97.9	1.6	3.1	0.6
SV*	**97.0**	97.0	2.8	0.0	0.0	**93.0**	93.4	6.4	0.4	0.0	**94.8**	95.6	4.3	0.8	0.0

* An asterisk denotes that the subject also participated in Experiment 4.

Table 3b: A comparison of the variability in localization accuracy among subjects in Experiment 4 (4 kHz cutoff frequency). For each stimulus condition the data are partitioned into corrects, adjusted corrects (i.e., sum of corrects and all reversals), errors, front-back reversals and back-front reversals. Data are expressed as means and are averaged across 4 sessions for each subject.

Subject	Stimulus														
	BB					HP 3 kHz					LP 3 kHz				
	Corrects	Adj Corr	Errors	F-B	B-F	**Corrects**	Adj Corr	Errors	F-B	B-F	**Corrects**	Adj Corr	Errors	F-B	B-F
AA*	**98.1**	99.4	0.5	0.7	0.6	**91.6**	99.8	0.2	0.1	8.1	**92.5**	99.2	0.7	1.8	4.9
BC*	**99.6**	99.6	0.4	0.0	0.0	**98.6**	98.6	1.4	0.0	0.0	**95.0**	98.8	1.2	0.2	3.6
DH*	**100.0**	100.0	0.0	0.0	0.0	**99.6**	99.6	0.4	0.0	0.0	**98.7**	99.7	0.4	0.5	0.5
EP*	**99.8**	99.8	0.2	0.0	0.0	**97.8**	97.8	2.2	0.0	0.0	**81.9**	100.0	0.0	4.9	13.2
JM*	**98.7**	99.2	0.4	0.0	0.5	**97.1**	97.5	1.8	0.0	0.4	**94.6**	98.2	1.8	0.2	3.4
MB	**99.9**	99.9	0.1	0.0	0.0	**97.2**	97.2	2.5	0.0	0.0	**86.7**	98.0	2.0	4.3	7.0
MG	**91.9**	91.9	7.8	0.0	0.0	**87.4**	87.8	12.1	0.4	0.0	**76.0**	91.8	8.2	3.1	12.7
MK	**99.8**	99.8	0.2	0.0	0.0	**97.1**	97.9	1.7	0.0	0.8	**62.0**	99.1	0.8	0.0	37.1
PC	**97.4**	97.4	2.2	0.0	0.0	**96.3**	96.3	3.4	0.0	0.0	**61.2**	98.6	1.2	0.0	37.4
PT	**98.6**	99.6	0.2	0.0	1.0	**82.8**	89.3	9.5	0.0	6.5	**82.9**	99.9	0.1	4.7	12.3
RA*	**94.7**	95.3	4.7	0.0	0.6	**92.5**	94.5	5.4	1.2	0.8	**92.2**	97.0	2.9	1.0	3.8
RT*	**98.3**	98.3	1.7	0.0	0.0	**97.4**	97.4	2.6	0.0	0.0	**94.1**	94.7	5.0	0.0	0.6
SV*	**98.9**	98.9	1.1	0.0	0.0	**93.4**	93.5	6.5	0.0	0.1	**95.6**	97.6	1.4	2.0	0.0

* An asterisk denotes that the subject also participated in Experiment 3.

Table 4: Localization performance in Experiments 1-4 (at least one each of VAS and Free-Field conditions per subject) at: (*a*) low-pass 3 kHz/4 kHz**, (*b*) high-pass 3 kHz/4 kHz**, and (*c*) broadband. For each experiment the data are partitioned into corrects, errors, and reversals. Data are averaged across 4 sessions for each subject.

(*a*)

Subject	VAS						Free-Field					
	Ex. 1 (quiet)			Ex. 2 (noise)			Ex. 3 (3 kHz cutoff frequency)			Ex. 4 (4 kHz cutoff frequency)		
	Corrects	Errors	Reversals	**Corrects**	Errors	Reversals	**Corrects**	Errors	Reversals	**Corrects**	Errors	Reversals
BC*				**43.5**	30.1	26.4	**85.9**	0.4	13.7	**95.0**	1.2	3.8
KS*	**43.7**	43.7	12.7	**28.9**	53.9	17.2	**91.7**	1.3	6.8			
RA	**60.1**	14.5	25.4	**52.7**	47.3	19.0	**89.3**	5.0	5.6	**92.2**	2.9	4.8
SD*	**34.4**	38.8	26.8				**34.4**	32.9	31.3			
SV	**41.2**	23.5	35.3	**42.8**	25.4	31.9	**94.8**	4.3	0.8	**95.6**	1.4	2.0

(*b*)

Subject	VAS						Free-Field					
	Ex. 1 (quiet)			Ex. 2 (noise)			Ex. 3 (3 kHz cutoff frequency)			Ex. 4 (4 kHz cutoff frequency)		
	Corrects	Errors	Reversals	**Corrects**	Errors	Reversals	**Corrects**	Errors	Reversals	**Corrects**	Errors	Reversals
BC*				**47.3**	36.6	16.0	**99.3**	0.7	0.0	**98.6**	1.4	0.0
KS*	**36.3**	46.5	17.2	**33.8**	47.2	18.9	**98.8**	1.0	0.0			
RA	**45.2**	22.2	32.6	**54.7**	45.3	17.2	**85.6**	9.0	5.4	**92.5**	5.4	2.0
SD*	**32.5**	41.7	25.8				**31.7**	32.8	33.5			
SV	**38.2**	25.4	36.4	**39.4**	60.6	29.1	**93.0**	6.4	0.4	**93.4**	6.5	0.1

(*c*)

Subject	VAS						Free-Field					
	Ex. 1 (quiet)			Ex. 2 (noise)			Ex. 3 (3 kHz cutoff frequency)			Ex. 4 (4 kHz cutoff frequency)		
	Corrects	Errors	Reversals	**Corrects**	Errors	Reversals	**Corrects**	Errors	Reversals	**Corrects**	Errors	Reversals
BC*				**49.1**	29.8	21.1	**99.8**	0.0	0.1	**99.6**	0.4	0.0
KS*	**44.4**	39.1	16.5	**33.7**	47.9	18.4	**99.4**	0.5	0.1			
RA	**63.4**	11.6	25.0	**55.5**	44.5	15.3	**93.5**	4.9	1.4	**94.7**	4.7	0.6
SD*	**36.5**	36.0	27.4				**36.5**	28.8	34.6			
SV	**42.4**	23.3	34.3	**40.9**	59.1	28.5	**97.0**	2.8	0.0	**98.9**	1.1	0.0

*An asterisk denotes that the subject did not participate in all experiments.

**The 3 kHz cutoff frequency applies to Experiments 1-3, and the 4 kHz cutoff frequency applies to Experiment 4.

DOCUMENT CONTROL DATA SHEET

1a. PERFORMING AGENCY DCIEM	2. SECURITY CLASSIFICATION UNCLASSIFIED
1b. PUBLISHING AGENCY DCIEM	

3. TITLE (U) The relative impact of generic head-related transfer functions and signal bandwidth on auditory localization: Implications for the design of three-dimensional audio displays
4. AUTHORS Arrabito, G. R.; Mendelson, J. R.

5. DATE OF PUBLICATION December 4 , 2000	6. NO. OF PAGES 77

7. DESCRIPTIVE NOTES
8. SPONSORING/MONITORING/CONTRACTING/TASKING AGENCY Sponsoring Agency: Monitoring Agency: Contracting Agency : Tasking Agency:

9. ORIGINATORS DOCUMENT NUMBER Technical Report 2000-067	10. CONTRACT GRANT AND/OR PROJECT NO. 6KD15	11. OTHER DOCUMENT NOS.

12. DOCUMENT RELEASABILITY Unlimited distribution
13. DOCUMENT ANNOUNCEMENT Unlimited

14. ABSTRACT

(U) Virtual auditory technology is being considered to cue armoured vehicle or air crew, via headphones of the communication system, to the spatial locations of potential lethal threats. Auditory localization in virtual auditory space (VAS) on the horizontal plane was investigated in this paper as a function of seven generic head-related transfer functions (i.e., digital filters for synthesizing the location of a sound in VAS), signal bandwidth (low-pass 3 kHz, high-pass 3 kHz and low-pass 14 kHz), and listening environment (quiet and in the presence of diffuse ambient Leopard tank noise). Testing was also conducted in the free-field which partially served to psychoacoustically validate the VAS conditions. The outcome of this preliminary study revealed that subject performance was better in free-field than in VAS. In the latter condition, subject performance was not significantly affected by type of generic head-related transfer function. Localization accuracy using the broadband stimulus was not significantly better than with the low-pass 3 kHz stimulus. Performance in the quiet condition was relatively better than in the noise condition. The implications of these results for implementation of a 3-D audio display into military environments and recommendations for future research are discussed.

15. KEYWORDS, DESCRIPTORS or IDENTIFIERS

(U) 3-D Audio; head-related transfer functions; sound localization; virtual auditory space: Implications for the design of three-deminsional audio displays